# Electric Contacts

## Theory and Application

By Ragnar Holm
Ph. D., hon. Dr. Techn.

in collaboration with Else Holm Ph. D.
both of St. Marys (Pa.)

FOURTH COMPLETELY REWRITTEN EDITION



Springer-Verlag New York Inc. 1967

/1J

Part I

# Stationary Contacts

## § 1. Introduction. A simplified summary of the theory of stationary electric contacts

An attempt to present the concepts of contact theory in such a sequence that none of them is used before being thoroughly defined would be unwieldy. For example, the fundamental concept of contact surface can not be treated in detail before the constriction resistance has been defined, since determinations of the conducting areas are based on the measurement of constriction resistances. Again, the concept of constriction resistance can not be thoroughly treated without reference to the contact areas. Therefore, it seems preferable to introduce, in a provisory manner, some important concepts in an opening paragraph, allowing a more elastic and agreeable exposition in those which follow. The introduction is given a quantitative character by means of some calculations, based upon an artificial model of the current flow lines which highly simplifies the treatment. In the later paragraphs the same model is used for the calculation of particular problems. Some of the concepts are defined solely in this introduction.

The term *electric contact* means a releasable junction between two conductors which is apt to carry electric current. These conductors may be called *contact members*, or simply *contacts*, when no misinterpretation is likely. The member from which the positive current enters the contact, is called *anode*; the other member is the *cathode*. When the contact members are separated by an insulating layer, it is conventional to speak of an *open* contact.

The force that presses the contact members together is the mechanical load or simply the *load*, $P$. If the contact members were infinitely hard, the load could not bring them to touch each other in more than three points. But since actual materials are deformable, the points become enlarged to small areas and simultaneously new contact points may set in. The sum of all these areas or *spots* ist the *load bearing area*, $A_b$, upon which the pressure, $p$, is finite. $A_b$ can be generated merely

by elastic deformation. But, because of the unevenness, the contact members, even though they may be nominally flat, actually touch each other in areas that are more or less plastically generated. They then satisfy Eq. (1.01)

$$P = \xi H A_b \tag{1.01}$$

where usually $0.2 < \xi < 1$ and $H$ is "the contact hardness"[1]. It is difficult to polish contact members so as to attain an average pressure as small as $0.05\,H$.

The area $A_b$ usually is much smaller than was supposed until about 40 years ago. In the case of nominally flat bodies placed on top of each other (as a brush on a ring), the whole covered area was often called the contact surface. It is more correct to call it the *apparent contact area*, as we shall do in the following. $A_b$ may be of a much smaller order of magnitude than the apparent contact area.

The expression *contact resistance* will often be used. This term was coined at a time when it was believed that the metallic contact surface itself accounted for the observed resistance. Actually, what is measured as contact resistance $R$, always implies or is a *constriction resistance* $R_c$, which is the consequence of the current flow being constricted through small conducting spots. Of course, the constriction extends into both contact members. An alien film in the contact may give rise to an additional resistance $R_f$.

One may theoretically expect that the discontinuity of the crystal lattice order in the contact will reflect electrons and produce a transition resistance. However, such resistances are extremely small and similar to the resistances of grain boundaries in a polycrystalline material. Measurements are unavailable. According to a rough computation these transition resistances per m² might be of the order of $\sigma = 10^{-8}\,\varrho\,\Omega\,\mathrm{m}^2$ where $\varrho\,\Omega$ m is the resistivity of the metal; cf. Eq. (3.04).

The surface $A_b$ usually is partly covered by insulating tarnish films and then only a fraction of $A_b$ has metallic or quasimetallic contact. A conducting contact area is called *quasimetallic* when it is covered with a thin film (20 Å or less) that is penetrable for electrons by means of the tunnel effect[2], and consequently produces a relatively small film resistance.

Summing up, we see that not only is the load bearing contact area very small, but also that only a fraction of it may be electrically conducting. In any case, the current lines of flow are bent together through narrow areas, causing an increase of resistance beyond the case of a fully conducting, apparent contact surface. This increase of resistance is the *constriction resistance*.

[1] Cf. § 1F. [2] See § 26.



The main part of the constriction resistance is localized in the immediate neighborhood of the small conducting spots. The probes measuring the contact voltage will always be at macroscopic distances from the contact surface. In most cases this distance covers the total range in which the lines of flow are deviated from a straight path by the constriction. Thus, the resistance measurement does not tell directly whether the measured resistance is essentially located within the contact surface or, more or less, has the character of a constriction resistance. We shall show that the constriction resistance can be calculated as a function of the conducting contact area, and when the measured resistance is greater than calculated for a known area, we conclude that the area is covered by a film that produces an additional resistance. Conventionally the expression *contact resistance* is preserved irrespective of whether there is merely a pure constriction resistance, or whether a film also contributes by a resistance at the conducting area.

We illustrate the constriction resistance by a calculation[1] under simplified conditions: assume a circular contact surface, same material in both contact members and perfect symmetry, i.e., no disturbance by any thermoelectric effects. The contact members are considered as semi-infinite. Even with these assumptions, the exact calculation is circumstantial (cf. § 4). However, by using the imaginary model illustrated in Fig. (1.02), we approach the real shape of the lines of flow with a much simpler calculation.

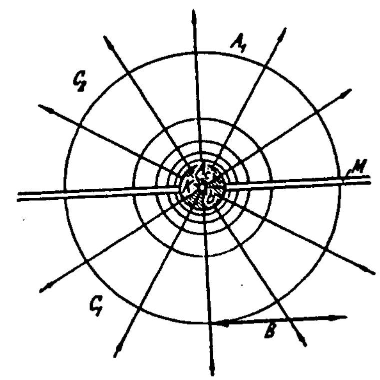

Fig. (1.02). Model to illustrate a current constriction with spherical symmetry. The resistance between two consecutive equipotential surfaces is 1/6 of the constriction resistance in one member

In the model the contact surface has been replaced by a sphere, $K$, of infinite conductivity, whose radius, $b$, is chosen slightly smaller than the radius, $a$, of the circular area; cf. Eq. (1.06). The lines of current flow start radially and are symmetrically distributed around $K$ so as to arrange the equipotential surfaces as hemispheres concentric to $K$.

Consider the constriction resistance in one contact member. The resistance $dR$ between the hemispheres with the radii $r$ and $r + dr$ is

$$dR = \frac{\varrho\,dr}{2\pi r^2}$$

[1] R. Holm [1].

if $\varrho$ is the resistivity of the contact material. We integrate through a hemisphere which passes through the point where the potential is tapped for measuring the contact *voltage*. The radius, $B$, of this hemisphere will be large compared with $b$. More distant parts add very little to the integral. We may therefore choose infinity as the upper limit and then obtain the constriction resistance in one contact member

$$R = \frac{\varrho}{2\pi}\int_b^\infty \frac{dr}{r^2} = \frac{\varrho}{2\pi b} \tag{1.03}$$

The total constriction resistance $R$ is twice as great, thus[1]

$$R = \frac{\varrho}{\pi b} \tag{1.04}$$

Eq. (1.04) is based on the artificial model (1.02) and therefore is approximate only. The correct value of the resistance belonging to a flat circular contact area with the radius $a$ is

$$R = \frac{\varrho}{2a} \tag{1.05}$$

as will be proven in § 4. Eqs. (1.04) and (1.05) define the same $R$ if

$$2a = \pi b \tag{1.06}$$

a relationship that plays a part in making deductions in later chapters.

The contact is heated by the current. If both contact members are of the same material, the highest temperature is localized in the contact surface and distant points are at correspondingly lower temperatures. Thus, if the supertemperature in the contact surface above the bulk of the contact members is $\Theta$, and the temperature coefficient of resistivity is $\alpha$, the constriction resistance will be increased by a factor somewhat less than $(1 + \alpha\Theta)$. This would be the factor in the case of a uniform temperature distribution. The factor for the resistance of the constriction, in which $\Theta$ is the maximum temperature, is approximately

$$\left(1 + \frac{2}{3}\alpha\Theta\right) \tag{1.07}$$

cf. § 16.

[1] It may be asked why the formula (1.05) contains the factor $1/a$ instead of $1/a^2$. The following simple consideration gives the answer. An electric resistance varies in proportion to a length and in inverse proportion to a cross-section. The dominating part of the constriction resistance is found in the neighborhood of the contact surface. The order of magnitude is for its length $a$ and for the cross-section $a^2$; thus the factor in question is $a/a^2 = 1/a$.



If we want to check this formula, the problem would be how to measure the temperature $\Theta$. Perhaps the first idea would be to try thermoelements fitted near the contact. This does not work because the elements never get close enough to the contact surface. The error would amount to the order of the measured value.

A simple, indirect method, that enables the determination of the supertemperature from the contact voltage exists because of the interesting relationship between the *supertemperature* $\Theta$ in the contact surface of a monometallic contact and the *contact voltage*, $U = RI$. This relationship is a consequence of the heat flowing along the same paths as the electric current, irrespective of the fact that the surroundings of the contact which insulate electrically do not insulate thermally. Because of the symmetry, there is no reason for heat transfer from one member to the other.

As is proved in § 13, said relationship for a monometallic contact with electric and thermal conductivities $1/\varrho$ and $\lambda$ respectively is

$$\int_0^\Theta \varrho\lambda\, d\vartheta = \frac{U^2}{8} \tag{1.08}$$

in the equilibrium state. Table (1.09) is calculated for copper. However, the table has a general validity for metals used in contacts, for according to the Wiedemann-Franz law, $\varrho\lambda$ is nearly the same for different metals[1].

Table (1.09). *Copper*

| | | softening | | melting | |
|---|---|---|---|---|---|
| $U$ | 0.03 | 0.12 | 0.3 | 0.41 | V |
| $\Theta$ | 10 | 190 | 700 | 1063 | centideg. |
| $1 + \frac{2}{3}\alpha\Theta$ | 1.04 | 1.5 | 2.8 | 3.8 | |

The table gives supertemperatures, $\Theta$, related to contact voltages, $U$, with the bulk at about 20 °C; and, below, the factor $[1 + 2/3\,(\alpha\Theta)]$. It gives the answer to the question raised concerning the method of determining the supertemperature $\Theta$ in a contact. We simply calculate it from the contact voltage using Eq (1.08).

It is particularly interesting that the voltage tells us whether the *melting point* is reached. The melting temperature is, of course, the highest temperature possible in a solid contact, and every attempt to

[1] Cf. § 111C.

increase the voltage beyond the melting voltage would result in the contact sinking together and forming a greater contact area that carries the current without further melting. If this happens, the voltage usually drops to a value somewhat below the melting voltage. Melting voltages of various materials are given in Table (XI.2).

The temperature can surpass the melting point only if the contact members are mechanically kept from approaching each other, for instance, in an opening contact. Then the boiling point of the metal may be reached. For copper, the boiling point corresponds to 0,8 V and for tungsten, 2.1 V. Boiling in the last moment of opening may immediately precede the ignition of an arc with an arc voltage of 10 to 15 V; since the arc demands a certain minimum gap, we conclude that boiling here appears like an explosion which at once produces a gap and fills it with overheated, ionized vapor.

In order to check the factor $[1 + 2/3\,(\alpha\Theta)]$ one must, of course, investigate a contact with a constant contact area. Usually, the contact area enlarges with increasing current because of softening of the metal by the heat. Therefore, in order to secure a constant area, one should begin the test with a high current and proceed to lower values. Or, better still, vary the current up and down and record *reversible variations* (see § 20). If such a reversible series of resistance records shows that the resistance increases with rising current (rising $RU$-characteristic, § 20), this would prove that the contact is metallic. Conversely, a falling $RU$-characteristic would indicate that the resistance belongs to either a semiconducting film, say a tarnish film, or to a thin, tunnel-conducting film. It happens that visibly tarnished contacts have a reversible $RU$-characteristic that rises with increasing current. Such a behavior indicates that the current flows mainly through (invisible) *metallic* contact spots which have been produced either by rupture of the film at contact make or by fritting.

A-fritting is an electric breakdown that occurs when the electric field reaches the order of $10^8$ V/m and may result in a metal bridge through the film[1] or even in a small conducting contact spot. There is a commonplace example of fritting. If you investigate an ordinary plug and socket contact by using a small emf, such as 1 V, you may frequently find it insulating. But, this feature is not noticeable in service since the ordinary line voltage is able to frit the contact.

We also distinguish B-fritting which leads to enlargement of the conducting areas at relatively low voltages across the conducting spot that is limited by a surrounding film. For details see § 27.

When the contact film is thin enough to be permeable to electrons

[1] Cf. § 27.



by means of the tunnel effect[1], usually no fritting occurs. A contact spot with such a film has been called quasi-metallic. The tunnel resistance may be negligible at high loads: but if the load is of the order of 0.01 N and the contact area correspondingly small, this resistance may surpass the constriction resistance and render measurements inconsistent. Many investigators have observed that it is necessary to submit such contacts to small vibrations before measuring in order to secure reproducible values. Holm called this action *aging* the contact[2].

The explanation of the aging seems to be as follows. When a clean metallic surface is exposed to air, it soon becomes covered by two layers of oxygen atoms. The outer layer is bound much weaker than the inner one but contributes to the tunnel resistance by a far greater amount. Later, these layers may develop into an oxide tarnish. It seems that aging results in the mechanical breakdown of the outer oxygen layer.

We have sketched problems of stationary contacts which constitute the very fundamentals of the theory of electric contacts. An extended treatment will be presented in Part I, while Parts (III and IV) will be devoted to problems of moving contacts. They concern: Part III sliding contacts, and Part IV make and break contacts, including the theory of the arc, the chief enemy of the switches. Although the electric conduction through sliding contacts ist physically the same as through stationary contacts, a significant difference may result from the fact that single contacts serve too short a time for reaching temperature equilibrium. This means that in sliding contacts Eq. (1.08) is no longer valid. Part II is devoted to thermal resistances.

## § 2. The contact surface

Concepts such as contact members, mechanical contact load $P$, load bearing area $A_b$, apparent contact area $A_a$ differentiated from the true contact area $A_c$, have been defined in the Introduction. The relation between the *load-bearing contact area*, the contact load, and the average pressure, $\bar{p}$, is

$$P = \bar{p}\,A_b \qquad (2.01)$$

The local pressure, $p$, may vary from point to point with elastic deformation in some spots and plastic deformation in others. In many

[1] Cf. § 26.
[2] In German it was called normieren; see R. Holm [29] p. 69; cf. § 20 B.

cases however, the entire area $A_b$ will yield plastically; for example, in contacts between two cylinders placed crosswise (cross rod contact) at an ample load. Then, the average pressure will be close to the hardness $H$[1]. Even in macroscopically flat contacts, $\bar{p}$ usually is not lower than about 0.2 of the hardness because a considerable fraction of the load-bearing elevations become plastically deformed; cf. § 7. It is to be noted that the hardness is not a strictly unambiguous concept, and the reader is referred to Appendix I and § 7 for detailed information on hardness and load-bearing area.

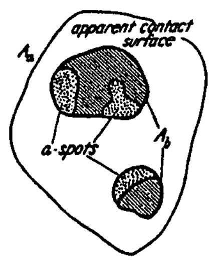

Fig. (2.02). Apparent contact surface $A_a$; load-bearing contact area $A_b$, containing insulating spots (shaded) and conducting spots, i. e., a-spots (dotted)

From the aspect of current conduction, the load-bearing area may consist of three different parts; cf. Fig. (2.02):

1. Portions with *metallic contact.* The current passes through them without perceptible transition resistance in the interface, just as it does between different crystallites in a compact metal; cf. Introduction.

2. *Quasi-metallic spots.* These are film-covered areas with films sufficiently thin to be easily permeable by the electron current by means of the tunnel effect, irrespective of the resistivity of the film material. Typical films of this kind are chemisorbed layers of oxygen atoms which, in air, are formed on any metal surface. This will be discussed in §§ 3, 6, 20, and 26.

3. Areas covered by relatively *thick films*; particularly, visible tarnish films (oxides, sulphides, etc.). As a rule, such areas are pratically insulating.

The short name *a-spot* for the conducting contact areas, referring to the radius $a$ of a circular contact area, is a widely accepted term.

Fig. (2.02) illustrates different kinds of contact surfaces. The ratios between these surfaces are diversified, and it may even happen that $A_a = A_b = A_c$. For example, if two *clean* metal cylinders, whose diameter is a few mm, are placed crosswise in contact with an ample load, 50 N, plastic deformation leads to a load-bearing area which is simultaneously the conducting area. This area is circular within the limits of the irregularities of the surfaces. It is a difficult problem to determine $A_b$ and $A_c$ in cases where these surfaces are small compared to $A_a$, as in the contact of a carbon brush on a collector ring. The present state of our knowledge concerning $A_b$ and $A_c$ is summarized in § 7.

The contact between *mercury and a solid metal* has particular features, since the deformation of the solid member is perfectly negligible

[1] As for hardness, see § I F.



and tarnish or chemisorbed films on it remain undamaged. Here the load-bearing area and the apparent contact area are equal, i.e., $A_a = A_b$. The contact between the film covered metal and mercury provides means for investigating the conduction of the films because we may regard $A_c = A_a$.

## § 3. The contact resistance. General theory

An introductory description of the basic ideas of constriction and contact resistance has been presented in § 1. We shall now enrich and widen the knowledge by forming the definitions on a broader basis, and it is felt that a typical example may illustrate their content without limiting the range of vision. Since thermoelectric forces are irrelevant for the concepts, we neglect them in the first instance and take them into consideration later. See § 18.

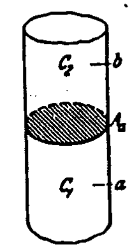

Fig. (3.01). The apparent contact surface $A_a$ is the common basesection of the cylinders which have real contact only in a portion of $A_a$

In Fig (3.01), the cylinders $C_1$ and $C_2$ are the contact members. They are placed on each other, base $A_{a1}$ against base $A_{a2}$, thus $A_a$ constituting the apparent contact surface. To begin with, we assume that the faces $A_{a1}$ and $A_{a2}$ are clean metallic but, due to their uneveness, have contact only in a small portion of $A_a$ namely in $A_c$. Because of the current flow being constricted through $A_c$, a "constriction resistance" appears; cf. § 1. It will be shown that this resistance not only depends on the size of the area $A$, but also on its shape. $A_c$ often consists of several spots, a-spots; cf. § 2. The voltage $U_{ab}$ may be measured between the points $a$ and $b$ when a current $I$ passes through the contact. Consequently the resistance between the equipotential surfaces containing $a$ and $b$ respectively is

$$R_{ab} = \frac{U_{ab}}{I}$$

We now imagine one single solid cylinder similar to the model except that the entire area $A_a$ is perfectly conducting so that the lines of current flow go straight through it. In this case let the resistance between the equipotential surfaces containing the points $a$ and $b$, the same as in the model, be $R^0_{ab}$. Then by definition, the *constriction resistance* and *constriction voltage* are

$$\left.\begin{aligned} R &= R_{ab} - R^0_{ab} \\ \text{and} \qquad U &= RI \end{aligned}\right\} \qquad (3.02)$$

In the case of a clean metal contact, i.e., with no film in the contact, the *contact resistance* $R$, is simply a *constriction resistance*. See § 1. If a film is present and both sides have different metals, the contact resistance $R$ is the sum of the constriction resistances $R_1$ and $R_2$ in the two members respectively and the *film resistance* $R_f$, thus

$$R = R_1 + R_2 + R_f \tag{3.03}$$

According to this definition, the contact resistance is not a transition resistance, as was believed earlier, but a surplus resistance in the body of the contact members, caused by the narrowness of the current paths through the a-spots, to which the resistance of a film on the a-spots may add. It is shown in § 5B that $R_1$ and $R_2$ are not quite independent of $R_f$.

As may be understood from the introduction, the order of magnitude of $R_1$ and $R_2$ is $\frac{\varrho_1}{na}$ and $\frac{\varrho_2}{na}$, where $a$ is an average linear dimension of the a-spots, $n$ their number, and $\varrho_1$ and $\varrho_2$ the resistivities of the members $C_1$ and $C_2$. A more accurate calculation of these resistances is given in § 4. The corresponding expression for $R_f$ is

$$R_f = \frac{\sigma}{A_c} \tag{3.04}$$

in the case of a conducting film uniformly distributed over the conducting area $A_c$, where $\sigma$ is the resistance across one cm² of that film. With $\sigma$ varying along the film, one applies

$$\frac{1}{R_f} = \int \frac{d A_c}{\sigma} \tag{3.05}$$

With $\varrho_f$ being the resistivity of the film material and $s$ being its thickness we have

$$\sigma = \varrho_f s \tag{3.06}$$

Through very thin films, the tunnel effect furnishes a current independent of $\varrho_f$ even if $\varrho_f$ is "infinite". For this case, $\sigma$ is defined and calculated in § 26. In later chapters the picture of the resistances $R_1$, $R_2$, $R_f$ will be completed by the study of details under different conditions.

Those regions within the contact members $C_1$ and $C_2$ (Fig. 3.07) where, owing to the smallness of the contact spots, the lines of current flow noticeably deviate from the straight course, are called *constriction regions* or simply *constrictions*. Within the constriction region, the potential gradient is relatively great but relatively small outside the constriction; in other words, $R^0_{ab}$ in Eq. (3.02) is relatively small, usually even negligible as compared to $R_{ab}$. As a consequence, an exact definition of the positions of the probes $a$ and $b$ ist not necessary. This



fact obviously contributed to the impression that the resistance between equipotential surfaces containing probes as $a$ and $b$ is located within the contact surface and to the designation of contact resistance for something that usually is a constriction resistance, localized in a very small but finite volume within the contact members.

Fig. (3.07) shows the system of equipotential surfaces and current flow lines when both members are of the same metal, and $A_c$ represents a single circular a-spot ($c$ means conducting) in the middle of $A_u$. In practice, the constriction may be regarded as limited in the bulk of the members by certain surfaces, as for example $A_e$ in the figure. Such surfaces we call *end-surfaces*. According to the figure, the distance of $A_e$ from the a-spot is of the same order of magnitude as the radius $r$ of the cylinder; and since $R^0_{ab}$ is of the order $\varrho B/\pi B^2 = \varrho/\pi B$, it is negligible compared with $R_{ab}$ if $B$ is very much greater than $a$. If the constriction volume compared with the volume $a^3$ is so great that $R^0_{ab}$ may be neglected, we speak of a *long constriction*; cf. Eqs. (4.09, (4.15) and (4.21).

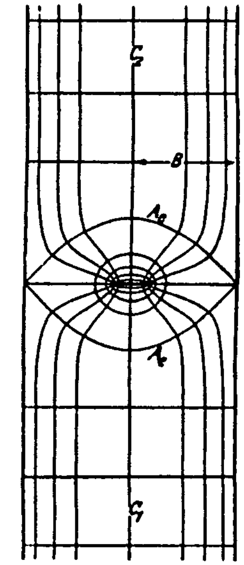

Fig. (3.07). Lines of current flow and equipotential surfaces of a current constriction

As already described, the smallness of $R^0_{ab}$ follows from the voltage gradient being small at distances from the a-spot which are much greater than the radius $a$ of the spot. When, however, a constriction is limited to a distance comparable with the radius $a$, $R^0_{ab}$ is no longer negligible and the *constriction* is called *short*.

Concerning the problem of determing the quantities $A_c$, $\sigma$, and the ratio $a/B$ which are implied in the formulas of this chapter, we refer to §§ 7 and 26.

## § 4. Calculation of constriction resistances with constant resistivity $\varrho$ in an isotropic material

We shall consider *long* constrictions; i.e., the conducting area $A_c$ is small compared with the dimensions of the "semi-infinite" contact members, and $R^0_{ab}$ in Eq. (3.02) is negligible. The constriction resistance $R$ is a function of the resistivity $\varrho$ of the material and the dimensions and shape of $A_c$. The general theory of the calculations of $R$ will be eluci-

With $l$ the average length and $A$ the average cross section of such an element, it is required that all elements satisfy $l/A = C$ where $C$ is a chosen constant. If $n$ such elementary tubes intersect an equipotential surface (the same tubes intersecting all equipotential surfaces) and if there are $m$ elements along each line of flow within the constriction, then the resistance of the distorted constriction is

$$R = \frac{m}{n} \varrho C \tag{5.18}$$

where $\varrho$ is the resistivity of the metal.

**Problem II.** A special distorted constriction has been treated with exactness by ROESS [1]. It concerns the contact between two cylinders in a central a-spot as illustrated in Fig. (3.07). The cylinder walls cause the system of flow lines to approach the type of parallel lines through a totally conducting $A_a$ corresponding to a vanishing constriction resistance. Let $R(a, B)$ be the constriction resistance, for one member, which for $a/B \to 0$ is $\varrho/4a$ and for $a = B$ is zero. For intermediate values

| | | | | | |
|---|---|---|---|---|---|
| of $a/B =$ | 0.1 | 0.3 | 0.5 | 0.6 | 1 |
| according to ROESS: $\frac{\varrho 2}{\pi B} \frac{1}{R(a, B)} =$ | 0.296 | 1.31 | 3.81 | 6.85 | $\infty$ |
| and with $R(a, B) = R_{eE}$ according to Eq. (5.08) with $n = 1$, $l = B$ and $A_r = \pi B^2$: | 0.296 | 1.29 | 3.79 | 6.82 | $\infty$ |

**Problem I.** Another kind of distortion appears when the current enters an infinite slab (thickness $h$, resistivity $\varrho$) through a circular contact area $A_1$ (radius $a$), with the bottom surface $A_2$ of the slab in perfect contact with an ideal conductor; cf. Fig. (5.19). Let $R(h, a)$ be the resistance in a slab characterized by $h$ and $a$. Approximate values $R(h, a)/R(\infty, a)$ can then be obtained from Fig. (4.28). The following table compares such calculations with measurements by FOXHALL and LEWIS [1]:

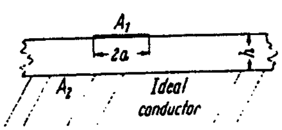


Fig. (5.19). Problem I. Current flow through a slab (thickness $h$) entering through the circular area $A_1$ (radius $a$), leaving through the infinite base $A_2$

| $h/a$ | 10 | 5 | 2 | 1 | 0.5 |
|---|---|---|---|---|---|
| $R(h, a)/R(\infty, a)$ accord. to Fig. (4.28) | 0.94 | 0.88 | 0.7 | 0.5 | 0.3 |
| " accord. to FOXHALL et al. | 0.96 | 0.90 | 0.80 | 0.64 | 0.43 |



## § 6. Introduction to thin films on contacts. Contact cleaning

Several following chapters are essentially devoted to perfectly clean contacts, but some short passages will refer to the behavior of thin, contact films. The main treatment of tarnish and other contact films appears in § 23–26.

The thin films now to be considered, particularly oxygen deposits, are of two kinds: physisorbed and chemisorbed.

*Physisorbed* particles (atoms or molecules) are bound to the metal by means of VAN DER WAALS forces with feeble bonds of the order of 0.05 eV. They are easily rubbed away at contact make, are to some degree volatile, and therefore can be pumped away at room temperature.

*Chemisorbed* atoms are bound to valences of the metal surface atoms by covalent bonds. In addition, the atoms may carry charges and be ionically bound. The bond strength is 1 to 8 eV. Two groups of chemisorbed atoms can appear on a surface, one stronger bound than the other. The difference is obviously caused by different sites on the pattern of the surface lattice. Chemisorbed films endure some friction and elevated temperatures without being removed from the metal surface. However, they are not thicker than about 10 Å and therefore easily penetrated by means of the tunnel effect (see § 26). As seen by the slopes of the associated solid and dashed curves in Fig. (8.01), the degree of influence of the tunnel effect on the contact resistance becomes smaller with increasing load $P$, and is for crossed rod contacts practically negligible at $P > 10$ N. But, the influence of these thin films on cold welding and friction can be considerable.

It evidently is important to define the concept of a *clean* metallic contact and to have reliable methods to accomplish the cleaning.

### *Clean metallic surface: definition*

With respect to contacts, a metal surface is "clean" when it does not contain any contaminant that perceivably influences contact tests: viz. measurement of electric contact resistance, cold weld or specific friction force. This requirement is harder to accomplish, the smaller the load. In any event, "clean" means that the surface shall not be covered with more than a small fraction of a monofilm and, of course, carry no alien insulating particles, as dust.

### *Methods to test the cleanness*

A. When the load bearing area $A_b$ is known and is a circle (example: cross rod contact under certain conditions at not too small $P$) essential

cleanness is stated if the measured contact resistance equals the constriction resistance associated with $A_b$ according to Eq. (4.15).

B. A friction coefficient of about $f = 1$ in air is an often used criterion of a fairly acceptable cleanness, cf. § 37.

C. At a cleanness similar to B, water wets the metal surface, viz. water drops spread on it.

Notice that oxygen from the air interferes so rapidly with clean surfaces that cleavage of graphite is much easier in air than in vacuum[1]; cf. § 39 A.

A *high degree of cleanness* can be achieved *in vacuum* alone.

Test methods are described in ROBERTS [*1*]. One sensitive phenomenon, for example, is the emission of electrons indicating whether the work function corresponds to the clean metal, and also secondary electron yield.

*Cleaning methods*

*Grease and lubricants* are removed with the aid of acetone, carbon-tetra-chloride and trichlor-ethylene (not as toxic as $CCl_4$). So much liquid must be used that the fatty solution becomes very diluted. Final cleaning is accomplished by washing in ethyl alcohol, and rinsing in tap water, or, better, in boiling distilled water.

Among the several recommended procedures for combined *mechanical* and *chemical cleaning* we cite[2]: Polish with 1/4 micron diamond powder or alumina; extract with benzene for four hours in a Soxhlet Extractor; soak for 15 minutes in hot chromic acid; rinse in destilled water; heat to 600 to 635 °C in helium atmosphere; use immediately to prevent contamination by the ambient atmosphere.

The cleanness of the surfaces and the density of the adsorbed films were determined by measuring the contact angle $\Theta$ of methylene iodide and water. For clean palladium $\Theta_{max} = 0$ to 3°. For organic monolayers, $\Theta_{max}$ was in the range of 37 to 93° depending upon the materials and coherence of the coverage.

Another fairly intricated method for cleaning is recommended by BLAKE [*1*].

For noble metals strong *etching* processes may be used. CHAIKIN [*1*] recommends for palladium, treating in cold aqua regia for 15 seconds, rinsing in 1 : 1 hydrochloric acid, and then rinsing in distilled water; this followed by a 15 minute treatment in sulfuric-chromic acid cleaning solution at 80 °C. Finally, after a thorough rinsing in double-distilled water, the contacts are dried in a desiccator containing magnesium perchlorate. The gold cleaning procedure may be similar except that warm (50 °C) aqua regia is used, the hydrochloric acid step may be eliminated, and the contact oven-dried at 110 °C.

[1] BRYANT et al. [*1*].

[2] Recommended by J. R. ANDERSON, Stanford Research Institute.



In many cases *cutting* the surface with a clean tool has been used. However, CHAIKIN has shown that an iron (or copper) tool can leave invisible traces of its metal on the treated surface. These may oxidize and cause trouble in micro contacts.

*Heating*[1] *in vacuum.* Atomically clean surfaces of high melting point metals can be generated by heating to temperatures, at which surface contaminants vaporize. It has been shown[2] that heating wolfram to 2200 °K for several seconds produces a clean surface, provided no contaminant is soluble in wolfram.

*Ultrasonic cleaning*[3]. Ultrasonic action provides efficient means to bring soluble and weakly adhering contaminants into solvent fluids. The action penetrates into crevices and pores. But, as a certain redeposit occurs, it is necessary to repeat the cleaning process in renewed clean solvents.

*Cleaning before assembly.* MAHLER [*1*] describes his great experience on making sealed-in contacts. He recommends: 1. to free every part from contaminants before its assembly; 2. to heat-stabilize organic components (if they are unavoidable) before they are used in the assembly; 3. to remove all impurities before the switch itself is placed in the housing; 4. to conduct all steps under aseptic conditions.

## § 7. The load bearing contact area as a function of load and elastic and plastic properties of the members

**A. Introduction.** The theory of indentation in § I treats the formation of calculable load bearing areas, $A_b$, for members with smooth surfaces and cases when $A_b$ is a circle or an ellipse, particularly the cases when the members are spheres, including a sphere and a plane, or crossed cylinders.

Imagine the indentation to be produced purely elastically at sufficiently small load. For a circular $A_b$, the radius is given by HERTZ's formula (I,1); for ellipses, formulas are given in ROARK [*1*]. These formulas are deduced for ideally smooth surfaces. Real surfaces have microscopic elevations and depressions, with the deep depressions remaining as voids in the contact. These usually much spread voids

[1] Cf. ROBERTS [*1*].

[2] Cf. HAGSTRUM et al. [*2*].

[3] Export rules for the operation are given in SCHROEDER [*1*] and L. K. JONES [*1*]. A description of the method and action is given in the article "Ultrasonics" in Encyclopedia of Electronics, Reinhold, New York 1962; cf. McCormick [*1*].

have a negligible influence on the contact resistance[1]. This is one reason why the formulas for smooth surfaces can widely be applied in this book.

When HERTZ formulas are applied to single asperity elevations, the concern can only be a rough computation since the curvature can not be well known. It is described in § I how plastic deformation begins and develops when the load on a contact with simple geometry increases and finally produces a "specific depth" of 0.03 or more. Then, the average pressure $\bar{p}$ attains $H$, the hardness. Interesting pictures of contact patterns with voids have been published by HÖFT [1] and GREENWOOD et al. [8].

B. Contact between *nominally flat contact members*, for example a carbon brush on a collector.

If the members were perfectly flat, there would be no constriction resistance. But actual surfaces have certain roughness (humps and grooves) and certain waviness (with macroscopic radii of curvature). At contact make, protuberances of any one member indent in the other member, and so the formation of contact areas resembles the formation of indentations described in § I.

At small average pressure, $\bar{p}$, the indentations may be formed elastically. With increasing $\bar{p}$, more and more indentations become plastically produced. Finally, nearly all indentations could have specific depths, $D$, of 0.03 or more. Then the load bearing area $A_b$ would satisfy

$$P = H A_b \tag{7.01}$$

with $H$ chosen so as to count for probable strain hardening.

However, this is an extreme and improbable situation. Actually, when some indentations deepen, other protuberances obtain the opportunity to make contact. These initially generate shallow, elastic indentations. The average pressure will be smaller than $H$, say

$$\bar{p} = \xi H \tag{7.02}$$

with $\xi < 1$. Hence

$$P = \xi H A_b \tag{7.03}$$

cf. Eq. (I,17).

Theoretically, any value of $\xi$ between 0 and 1 is possible; but according to measurements, values between 0.1 and 0.3 are most frequent for apparent contact pressures which are not too small, say for

$$p = \frac{\text{load } P}{\text{apparent area } A_a} > 10^3\,\mathrm{N/m^2} = .145\ \text{psi}$$

R. HOLM[2] has reported $\xi$-values as small as 0.02, obtained in a carbon-carbon contact after extended polishing of the members against

[1] Cf. Fig. (5.16).
[2] R. HOLM [37] p. 35; cf. R. HOLM [1] p. 323.



each other, steadily back and forth in the same straight path. Recently, CLAUSING and CHAO [1] attained $\xi < 0.02$ between extremely smooth contact members of stainless steel which had a curvature radius of between 50 and 100 m. The load was about 1000 N.

In § 36 B the explanation of COULOMB's law of friction is discussed. Thereby the question arises whether $\xi$ can be independent of $P$. As to this particular problem we point to the fact that $\xi$ is proportional to the average pressure $\bar{p}$ (see I,17 and the remark connected with this equation). Therefore, $\xi$ is a function of the specific depth $D$ of the indentations constituting the load bearing area. Thus, a constant $\xi$ means a constant average specific depth $D$. Since nature evidently presents cases of $\bar{D}$ = constant, it is important to investigate relevant conditions.

According to Eq. (I,17) a constant $\xi$ is equivalent to $A_b \propto P$. GREENWOOD and WILLIAMSON[1] have investigated a mathematical model of a very general type that satisfies $A_b \propto P$. They characterize the unevenness of a surface by asperities each of them represented by its height, $z$, and the radius, $r$, of curvature of its top. They show by profilometric measurements that usually a Gaussian distribution of $z$ is essentially realized. The distribution of $r$ is skew.

The authors assume a contact between nominally flat members (without macroscopic waviness) with the asperities so far apart that they deform independently of each other. When the higher asperities deform, lower ones touch, and a variety of contact areas with different sizes is produced. It is shown that probably both the number of spots and the total contact area are fairly proportional to the load $P$, whereas the density of the spots is proportional to the apparent pressure $P/A_a$, where $A_a$ is the covered (apparent) area.

The authors introduce the concept "plasticity index", $\Psi$, that we slightly modify to

$$\Psi = \frac{E}{H}\sqrt{\frac{\sigma}{r}}$$

with $\sigma$ = standard deviation of $z$, and

$$\frac{2}{E} = \frac{1}{E_1} + \frac{1}{E_2} \quad \text{and} \quad \frac{2}{H} = \frac{1}{H_1} + \frac{1}{H_2}$$

where $E$ = YOUNG's modulus of elasticity and $H$ = hardness. Indices refer to the member 1 and 2. When $\Psi < 1$ and the apparent pressure $< 10^3\,\mathrm{N/m^2}$, all deformations are practically elastic.

[1] J. A. GREENWOOD et al. [1], [5], [7], [8], where earlier contributions, particularly that of ARCHARD [2] are discussed.

## § 8. The relation between contact load and resistance, particularly at moderate and high load

**A. Introduction with description of Fig. (8.01).** In a clean contact between sufficiently smooth members in the shape of two balls, a ball and a plate, or two equal cylinders that are crossed, the load bearing area $A_b = A_e$ conforms with HERTZ equation (I,1) if the indentation is purely elastic. When the deformation becomes plastic, $A_b$ will correspond to Eq. (7.03), and to Eq. (7.01) when the indentation attains a specific depth above 0.03 (totally plastically deformed). In all these events, and between isotropic materials, the contact surface is circular (elliptic between crossed rods of different diameters); and if its linear dimensions are less than 1/20 the dimensions of the contact members, the constriction is long and its resistance can be calculated according to Eq. (4.14).

Imagine curved members with perfectly smooth surfaces, say, two equal cylinders with radius $r$ pressed cross wise into contact by a load $P$ producing purely elastic deformation. The contact area is a circle, $A_b$. With a greater $r$, a greater $A_b$ is formed by the same $P$. Theoretically, with the curvature radius infinitely increasing, $A_b$ would also infinitely grow. But, a practical limit is defined by the always existing waviness of the surfaces. This implies, that with nominally flat members, $\xi$ of Eq. (7.03) decreases with decreasing waviness of the surfaces, and, if the waviness ist not known, $\xi$ is uncertain between wide limits, say between 0.02 and 1; cf. § 7B.

Nevertheless, bolted junctions, for instance between bus bars and many other contacts between apparently flat members, show resistances that scatter astonishingly little at given load $P$ when the surfaces are clean. The faces of the members may have been fairly flat before being bolted, but the load usually deforms them. This has the effect of concentrating the contact spots to a rather small area. This means an approach to rod contacts of a relatively large rod diameter, implying a fairly unequivocal relation between $P$ and $R$. That is, a representing curve can be drawn for any metal in Fig. (8.01), around which observed points gather with scarcely any deviations towards higher $R$ but with considerable deviations below the curve in cases of very thick members (resistant against deformation) with extremely small waviness. On Fig. (8.01) the curves marked Cu-plates and Ni-plates illustrate this relation between $R$ and $P$. A more thorough treatment of nominally flat contacts is given in § 35 on thermal contacts.

Fig. (8.01) has been designed[1] to meet the practical demand of an easily read graphical illustration of the relation between $P$ and $R$ in

[1] Fig. (8.01) essentially constitutes an extract from several similar diagrams in R. HOLM [4].



metallic contacts under conditions given in the introduction of this chapter. The figure refers to: A) contacts between crossed cylindrical rods (labeled rods); B) junctions between nominally flat members, for

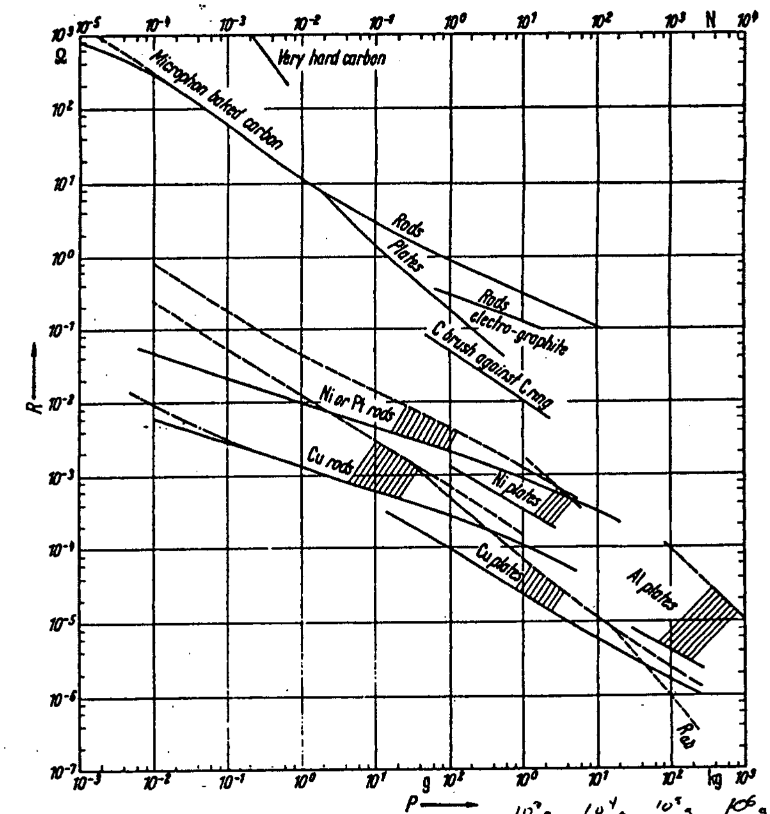


Fig. (8.01). Virgin contact resistances $R$ plotted against the contact load $P$. Solid lines are for clean contacts. In air the metals soon become covered by a monolayer of oxygen. The increased resistances are represented by the dashed lines, associated with solid lines by shading. Practical contacts have thicker alien films. Their curves have a steeper slope and often lie higher above the dashed lines as the latter above the solid lines. The measurements for the carbon have been carried out in air, but also hold for vacuum, except for their dashed part

example contacts between bolted bars with rectangular cross section (labeled plates).

For any material, four curves are drawn, namely: 1. solid for clean crossed rods; 2. dashed and connected with its associated curve No. 1 by shading for the same rods when covered by a chemisorbed oxygen film; 3. solid for clean nominally flat members; 4. dashed for the same nominally

flat contacts when covered with a chemisorbed oxygen layer, the association with the corresponding solid curve again being indicated by shading.

Every observation was made with a new contact. The dashed curves are entirely based on measurements on contacts which were closed after being exposed to air for some minutes (up to one hour) after a thorough cleaning. Under these circumstances, we can expect a chemisorbed oxygen layer to cover the faces. This seems to be confirmed because the deviation, $R_f$, from the associated solid curves can be correctly calculated from Eq. (8.02) assuming reasonable values for the tunnel resistivity $\sigma$ of the film

$$R_f = \frac{\sigma}{A_a} \tag{8.02}$$

The solid curves for *nominally flat contacts* (plates) as well as all curves for carbon are also based on measurements. However, the solid curves for metallic rods have been calculated and checked in vacuum by measurements in only a few points.

The basic data for the calculations are summarized in Table (8.09) which is followed by an example. Notice that Fig. (8.01) refers to circumstances that are idealized in certain respects. Particularly, the solid curves concern perfectly clean and very smooth surfaces. In practice, deviations are caused by surface contamination and roughness. General rules for these changes can not be given, but it is important to know the optimum that can be attained. This is presented in Fig. (8.01).

B. Crossed rod contacts. Preceding the discussion of the results, we first indicate by means of Figs. (8.03) to (8.06) a suitable method for measuring contact resistances which is the same method that was used, in 1928, by R. Holm [4]. Fig. (8.03) illustrates the wiring diagram, while Figs. (8.04) and (8.05) show an appropriate construction of the holders for the contacts. A circular contact area $A_c = A_b$ is obtained under conditions described in the introduction of this chapter. Fig. (8.06) pictures the equipotential surfaces in one of the cylinders. In the vicinity of the contact area, represented by a point in the figure, the equipotential surfaces are nearly concentric ellipsoids of the type shown in Fig. (4.16).

The surface $A_c$ is considered as end surface of the constriction. Since it intersects the end of the cylinder, the voltage measured according to diagram (8.03) will be the voltage between the end surfaces, and consequently be what is called the contact voltage, $U$. With $I$ being the current, the ratio $U/I$ is the contact resistance. $R$, to be measured with sufficiently low $U$ to avoid heating of the constriction. If the radius of the cylinders is sufficiently large, say more than 20 times larger than that of the contact surface, (cf. Fig. [4.28]), the constriction is long and its resistance can be calculated according to Eq. (4.15).



***Results of the resistance measurements with crossed rod contacts.*** To begin with, we consider a cylinder material of a moderate hardness, for example copper. The diameter may be $2r = 5$ mm and we suppose the contact load to be above 100 N. Then the pressure $p$ will attain the plas-

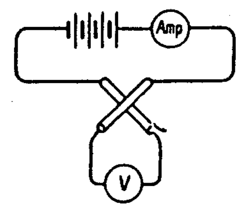


Fig. (8.03). Wiring diagramm for measurement of contact resistance between crossed rods

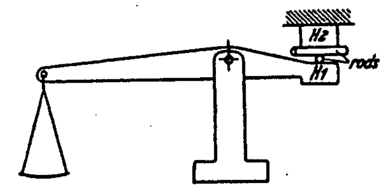


Fig. (8.04). Balance used for resistance measurements on contacts between crossed rods. Each rod is kept in position by a holder $H_1$ or $H_2$ in such a manner that bending is avoided. $H_2$ is fixed, $H_1$ is connected to the balance lever

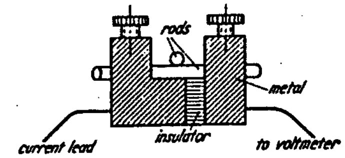


Fig. (8.05). Form of the holder $H_1$ and $H_2$, used for the device shown in Fig. (8.04)

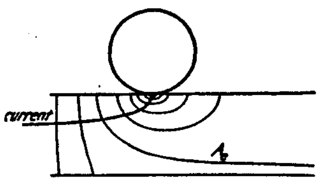


Fig. (8.06). Equipotential surfaces in the environment of a contact between crossed rods

ticity limit in nearly the whole contact surface, making the impression circular. The constriction will be fairly long and Eq. (4.15) applicable with good approximation.

We then consider a smaller load, for example $P = 0.1$ N, making it possible for merely elastic deformation to produce the contact surface; note that this is true only for sufficiently smooth rod surfaces. The contact area will again be a circle, and its radius $a$ is given by Eq. (I,1). Combining (I,1) with (4.15) and (7.01) with (4.15) we easily find in the case of a small $P$ and purely elastic deformation

$$R \propto P^{-1/3} \tag{8.07}$$

and in the case of a large $P$ and purely plastic deformation

$$R \propto P^{-1/2} \tag{8.08}$$

Both equations are true, of course, only for clean metal surfaces and, beyond this, in the case of (8.07) for sufficiently smooth surfaces.

Although surfaces of real bodies are never perfectly smooth but affected with a certain asperity[1], the contact surface at large $P$ will be

[1] Amplitudes of the asperities 0.1 to 0.01 mm, cf. Schmaltz [1], and p. 32

essentially coherent because the protuberances of one face are pressed into the counter face[1]. The contact resistances, under these circumstances, are well represented by Eqs. (4.15) or (4.21). As for the influence of persisting grooves within the contact surface on the resistance, cf. Fig. (5.13).

We now suppose the contact load to be sufficiently small to produce with perfectly smooth members a contact area $A$ by merely elastic yielding, where $A$ is the so-called HERTZ area corresponding to Eq. (I,1) with $r$ being the macroscopic curvature. However, with asperity being superposed on $A$, it may happen that only discrete a-spots are generated, mainly plastically as discussed in § 7 with Fig. (7.15a). Then, the contact resistance is given by Fig. (5.13).

It is obvious that this case may be realized with microphone carbon contacts at $P < 0{,}2$ N. Here the contact resistance is largely independent of the curvature of the contact members, i.e., independent of the size of the HERTZ area, so that even the curve for nominally flat members coincides with the curve for rods in Fig. (8.01). The explanation[2] for this fact is as follows. The contacts in question have a HERTZ area with discrete contact spots. For a very small load, $P$, there may be only three spots. With rising $P$ the size of these spots increases only slightly. At the same time, new small a-spots are added with the result that the average radius of an a-spot, thus the resistance per spot, remains essentially constant. Such a resistance phenomenon is fairly independent of the curvature of the members as long as the a-spots remain as discrete spots on the HERTZ area in the fashion of Fig. (7.15a). In other words, the resistance is independent of the HERTZ area, and as a consequence, independent of the curvature of the members.

C. Explanation of the dashed lines in Fig. (8.01). As already mentioned in this paragraph, the deviation of the dashed lines from the

Table (8.09). *Data for diagram (8.01)*

$r$ = radius of rod, $E$ = YOUNG's modulus, $H$ = hardness measured with the specific depth $D \approx 0.04$, $\varrho$ = resistivity, and $\sigma$ = tunnel resistivity

| Material | $2r$ cm | $E$ $10^{10}$ N/m² | $H$ $10^{8}$ N/m² | $\varrho$ $10^{-8}\,\Omega$ m | $\sigma$ $10^{-12}\,\Omega$ m² |
|---|---|---|---|---|---|
| Al | — | 6 | 2.4 | 2.8 | — |
| Cu | 0.5 | 11 | 5 | 1.76 | 1.5 |
| Pt | 0.08 | 15 | 6.5 | 13 | — |
| Ni | 0.3 | 20 | 14 | 9 | 2 to 8 |
| Graphite | 0.5 | 0.8 | 1.4 | 4300 | — |
| baked carbon | 0.3 to 0.6 | 0.5 | 3.5 | 7000 | 7 |

[1] See, for instance, R. HOLM [28] Fig. 2, and I. MING FENG [1] particularly Figs. (5) and (8).

[2] R. HOLM [1] § 7 and R. HOLM [4] p. 243.



course of the associated solid lines is explained by the existence of a film penetrable for tunneling electrons. Table (8.09) contains basic data for Fig. (8.01).

We calculate a point on the copper rod curve obtained with very smooth and clean surfaces. At $P = 0.1$ N we are in the region where the yielding is elastic. Using the data of Table (8.09) and Eqs. (I,2) as well as (4.15), i.e., long constriction, we obtain

$$a = 1.11 \sqrt[3]{\frac{0.1 \cdot 2.5 \cdot 10^{-3}}{11 \cdot 10^{10}}} = 1.46 \cdot 10^{-5}\,\mathrm{m}$$

and

$$R = \frac{\varrho}{2a} = \frac{1.76 \cdot 10^{-8}}{2.92 \cdot 10^{-5}} = 6 \cdot 10^{-4}\,\Omega$$

$R$ can be read on the solid line. The indentation is elastic since the average pressure $\overline{p} = P/\pi a^2 = 1.5 \cdot 10^8\,\mathrm{N/m^2} < H/3$ is too small to produce a plastic indentation, cf. § I. The error resulting from assuming a long constriction is $< 1\%$ in the actual case of $\frac{r}{a} = \frac{0.25}{1.46 \cdot 10^{-3}} = 170$, i.e., $\sqrt{\mu} = 170\,a$. The additive resistance, $\Delta R$ (leading to the dashed line), which we assume to be the resistance of a thin uniform film, obtained from Eq. (8.02), is

$$\Delta R = \frac{\sigma}{\pi a^2} = 2.24 \cdot 10^{-3}\,\Omega$$

Thus, the total resistance corresponding to $P = 0.1$ N is found to be

$$R = 0.6 \cdot 10^{-3} + 2.24 \cdot 10^{-3} = 2.84 \cdot 10^{-3}\,\Omega$$

differing very little from $2{,}9 \cdot 10^{-3}$ as given by the dashed line.

The slope of the solid line in the region $P = 0.1$ N is $-1/3$, according to Eq. (8.07). Evidently, $R$ is dependent on $r$ in the case of elastic deformation between relatively smooth surfaces according to Eq. (I,2), but independent of $r$ when the deformation is plastic.

Applying $P = 50$ N would bring about plastic deformation and $R$ has to be calculated according to Eq. (7.01). The slope then is $-1/2$, according to Eq. (8.08). The portions of the graph with the slopes $-1/3$ and $-1/2$ are connected by a slightly curved line.

The influence of the waviness has not been taken into account when drawing the solid lines, otherwise they would have turned upwards a little at their left end. But, the measured dashed lines show this tendency.

Notice that the dashed curves have a greater slope than the solid ones. In the event of films thicker than those for Fig. (8.01), the slope of the resistance lines approaches proportionality to $P^{-1}$. The reason is that the dominating film resistance varies as $A_c^{-1}$ according to Eq. (8.02), and $A_c$ is nearly proportional to $P$.

**D. Diversified resistance measurements.** Diagrams with $R$ plotted against $P$ ($R$-$P$-characteristics), measured under different conditions, have been given by many investigators. We cite R. HOLM [4] § 12, CONTIUS [1], MÜLLER-HILLEBRAND [1], KAPPLER et al. [1], BOWDEN and TABOR [2], SHOBERT [1], COOKS [1], FUKUROI and MUTO [1], and FAIRWEATHER [1]. FAIRWEATHER carefully avoided artificial aging. Nevertheless, his results agree very well with Fig. (8.01) which means that aging has little influence on clean contacts if $P$ is not too small, say $> 1$ N. According to KAPPLER et al., silver and gold, freshly scraped, did not show any alien film resistance, and the results on crossed-rod resistances coincide very closely with the solid copper rod curve in Fig. (8.01); but after the members were exposed to air for some weeks, the contacts became covered by a film with $\sigma = 5 \cdot 10^{-13}\,\Omega\,m^2$. Also, on platinum such a film appears minutes after scraping, even though it is somewhat non-coherent. On nickel, a film with $\sigma = 2 \cdot 10^{-12}\,\Omega\,m^2$ was observed.

COOKS determined average values of $\sigma$ for various metals in air after practical cleaning. He found, for example, with both members of nickel or wolfram, $\sigma$ of the order of $10^{-10}\,\Omega\,m^2$. With wolfram against gold, $\sigma$ was of the order of $10^{-9}\,\Omega\,m^2$. The gold very likely did less damage to the film.

MILLIAN and RIEDER [1], COMPTON and BAKER [1], cf. ANGUS [1] published measurements on resistances in cross-rod contacts which had been cleaned in different ways and were observed in air. The significant conclusions can be so expressed: On freshly cleaned contacts $\sigma$ is of the order of $10^{-12}\,\Omega\,m^2$. Even after a month, and with $P > 10$ N, one finds $\sigma \approx 2 \cdot 10^{-12}$ on silver and electropolished wolfram. But on copper, $\sigma$ grows to more than $10^{-11}$ after a month and $10^{-9}$ after 3 months, all with the unit $\Omega\,m^2$.

An interesting feature of the rod- and flat contact curves for carbon, as has already been described, is that they coincide for $P < 0.02$ N. This is explained above as a result of the asperity of the faces, which shapes the contact area as a group of discrete a-spots, fairly alike, whether the members are rod-shaped or flat. Alien films exist, very likely chemisorbed oxygen, but they interfere imperceptibly at $P > 2 \cdot 10^{-4}$ N, since their resistance is small compared with the considerable constriction resistance in carbon. Their role at smaller loads will be discussed in § 9.

*Short, distorted constrictions.* Until now, we have calculated with *long constrictions* and assumed the thickness of the rods or bars to be great compared with the diameter of the individual contact spots, and the spots to lie sufficiently apart to prevent distortion of the constrictions. An interesting exception is noted by the dotted line in the



lower right of Fig. (8.01). The measurements in question were made with crossed bars pictured in Fig. (8.10). The bars had been greased, brushed by means of a steel wire brush, and wiped clean without completely removing the last layer of grease, and they were then immediately clamped together. This procedure produced fairly clean contacts.

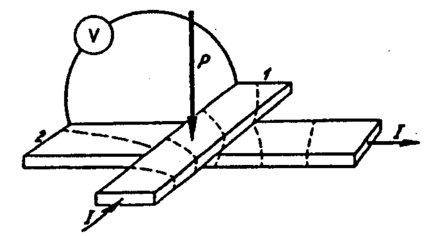

Fig. (8.10). Cross bar contact with negative $R^0_{ab}$

The bars were 15 mm broad and 3 mm thick. It is obvious that the constriction was not long and the tangential resistance in the bars became comparable with the constriction resistance. This can cause the quantity $R^0_{ab}$, defined in § 3, to assume negative values, and in fact negative quantities for $R_{ab}$ were measured with $P = 10^4$ N. To understand this, we consider the extreme case in which $P$ is large enough to bring the whole covered surface into actual contact. Then the equipotential surfaces will intersect the contact as indicated in Fig. (8.10). We see that the potential tapped at point 2 will be that of an earlier equipotential surface than is tapped at point 1. Thus, the voltmeter measures a voltage with reversed polarity.

**E. Use of Diagram (8.01) in practice.** The diagram is very helpful for determining: 1. the contact resistance corresponding to a certain load, or 2. the load needed if the resistance is prescribed. The causes for deviation from the curves in Fig. (8.01) are often found to be in a faulty wiring or in films thicker than those met in the dashed curves.

It is of no great disadvantage that only some few metals are represented, since the curves for all metals have a similar inclination. Therefore, in most cases it is sufficient to calculate one point in the elastic and one point in the plastic region, and then draw the characteristic parallel to the plotted ones. The formulas for the calculation are given in §§ 4, 5, 7, 8, and I.

For a rough computation of the resistance in a contact with $0.1 < P < 100$ N, practically clean members with resistivities $\varrho_1$ and $\varrho_2$ and the contact hardness, $H$, of the softer member, one may use the following rule of thumb:

$$R = \frac{\varrho_1 + \varrho_2}{2} \sqrt{\frac{H}{P}} \qquad (8.11)$$

**F. Practically clean contacts, preloaded with a high $P$**[1]**.** Comparison of the resistance $R$ at the light loads $P$ in Fig. (8.12) with that of the

[1] R. HOLM et al. [8] p. 61. Fig. 14a.

nickel curves in Fig. (8.01) at the same loads shows that the films on the practically clean nickel rods were relatively thick, but still tunnel conducting. Submitting these contacts to great loads reduces $R$ quicker than according to (8.01). Varying $P$ between high and low (without vibrations) results in a constant[1] $R$ for a pressure interval between 1.9 and 0.25 N. This means that metallic spots[2] have been produced (most probably on the elevations) in which the members were welded with a strength that could withstand the elastic counterforce during unloading. The average pressure attained about half the hardness of the nickel, which makes it very likely that complete plastic deformation was produced in the contact at the top of elevations, disrupting the continuity of the film in some spots.

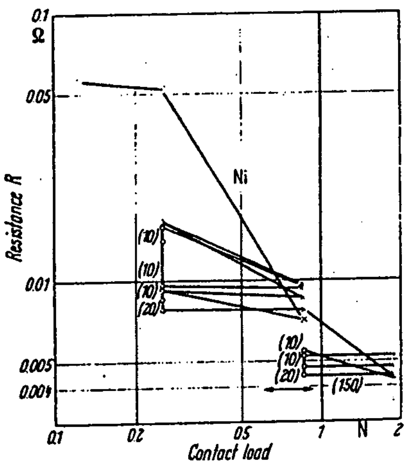


Fig. (8.12). Cycling of the load $P$ (50 to 150 times) on a crossed nickel rod contact; freshly cleaned rods of 0.3 cm diameter; exposed to air two days

## § 9. Contact resistance on freshly cleaned rods in air at very small contact loads

**A. Observations on gold and silver.** Instructive measurements were made in F. L.[3] with the aid of a modified moving-coil instrument. The pointer, provided with a very good bearing, carried one contact member, the other being stationary, see Fig. (9.01). The contact load, or sometimes tension was generated by means of the current through the coil. Fig. (9.02) shows observations on crossed gold and silver rods. The conductance $1/R$ is plotted against the load $P$. The freshly cleaned rods were washed with soap and water and finally with alcohol. Sometimes

[1] R. Holm [8] p. 61. Aluminum contacts behave differently due to the rapid oxidation of clean surfaces. A metallic spot which was welded oxidizes very quickly when the contact is broken and the spot exposed to air. When the contact is remade this spot no longer exhibits adherence.

[2] Holm [26] pp. 334 and 335.

[3] R. Holm [4] § 13 and R. Holm [30] Fig. (16.02) for which the measurements were also made in F. L.



they were also scraped with a clean tool, but there always was sufficient time to allow a deposit of the normal (chemisorbed) oxygen layers on silver. Without the film (perfectly clean) contacts would yield a conductance at $P = 10^{-4}$ N several times larger than was observed.

An interesting feature of the silver curves in Fig. (9.02) are the sudden changes of the conductance to higher values. We contribute these jumps to the intricate nature of the formation of a metallic spot as follows. We assume one rod to have contact with a single protuberance. First, the contact is formed merely elastically and the conductance through the film is due to the tunnel effect. With increasing load the load carrying hump breaks down, disrupting the film and producing a clean a-spot. This is evidenced because after such a yield the contact shows a cold weld during unloading: cf. Fig. (9.02). Gold contacts show less jumps; but whenever a jump occurs the contact exhibits a cold weld. The metallic spots produced are, of course, very small.

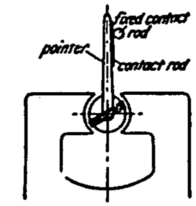


Fig. (9.01). Device for measuring contact resistances at very small loads $P$

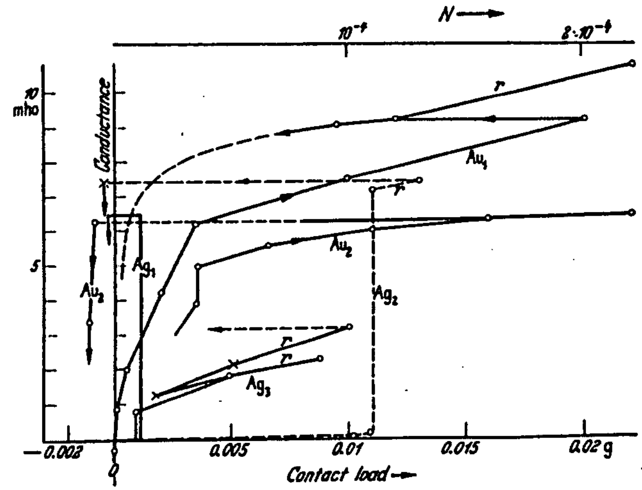


Fig. (9.02). Conductance $1/R$ of gold and silver contacts plotted against the load $P$ in g = 0.01 N. Hardness $H$ (silver) = $7 \cdot 10^{8}$ N/m². The reversible branches which have been run several times, are marked with $r$

At further increasing load, there may be new rupturing of the film on the same hump due to new plastic yielding which would result in enlargement of the already existing metallic spot. Or, other humps may collapse too, and new spots may be generated. Very likely both phe-

nomena will happen during one jump. It may even be that the effect of sudden enlargements of a-spots is involved to a greater extent than the generation of new spots with the consequence that the number of jumps does not equal the number of a-spots. The suddenness of the jumps indicates that the yielding protuberances were overloaded, i.e., loaded with a higher pressure than the hardness $H$. FAIRWEATHER [1] has observed similar jumps.

The *reversible branches* (symbol $r$) evidently are due to elastic variation of that portion of the contact area which is quasimetallic, whereas the produced metallic spots may remain constant due to adherence.

The adherence or cold welding of contacts is often troublesome in service. In Fig. (9.02), a minor adherence is indicated as an example in the curve $Au_2$. After having attained the load $2.2 \cdot 10^{-4}$N, $P$ was decreased, but the resistance did not increase. The contact area kept adhering without altering its size until a tension of $8 \cdot 10^{-5}$N separated the contact members. The platinum contacts did not adhere to any large extent and did not offer as good and uniform a conduction at small loads as gold contacts. Nickel contacts at small loads could compete with platinum contacts were it not for their magnetic properties which produce a particular kind of adherence; see p. 59.

The measured adherence indicates that clean metal surfaces stick to each other as if welded together, cf. § 28. If the contact area remains unchanged on removal of the load, one can expect the adherence force per m² to be equal to the tensile strength $Z \approx 0{,}3\,H$ of the metal in question. Then at contact break the tension would reach 0.3 of the initial pressure. But even with perfectly clean surfaces, one measures less because elastic forces help to lift the contact; cf. § 28.

The kind of adherence described above must not be confused with another kind, very common in practice, produced by liquid contaminations on the contact surface. This kind is prevented by cleaning.

**B. Observations on carbon contacts.** The curves for carbon contacts, in Fig. (9.03), show the influence of alien films because they do not begin to rise at the zero load point. The jumps are strongly marked. We explain them as caused by a sudden generation of clean a-spots by disrupting of the surface films when the underlying carbon yields plastically on suitably shaped load bearing humps, or the enlargement of existing a-spots or both. After any sudden change, a gradually rising branch appears. It is reversible (symbol $r$) and consequently belonging to merely elastic deformations with negligible adherence. This is the reason why the reversibles have a smaller slope than the curves, on an average.

Fig. (9.04) reproduces the reversibles with other co-ordinates and compares them with curves from Fig. 21 in GOUCHER [1]. These were also obtained with iterated load variations.



Attempts by GOUCHER[1] to explain the effect in microphone contacts by assuming hemispherical humps on the grain surfaces do not lead to increased understanding. The shape of the humps is obviously other than hemispherical.

As mentioned above, we expect the resistance of the chemisorbed film on carbon to be measurable at $P < 2 \cdot 10^{-4}$N. In fact, observations of WRIGHT and MARSHALL [1] show this effect. They investigated the contact between two carbon filaments from incandescant lamps, both after degassing at 1700 °C in vacuum and after exposing them to air. The measurements give, at $P = 10^{-4}$N: $\varrho/2a = 80\,\Omega$ and $\sigma/\pi a^2 = (125 - 80)\,\Omega = 45\,\Omega$. Substituting the resistivity $\varrho = 3.5 \cdot 10^{-5}\,\Omega$m of the carbon used we calculate $\sigma = 6.7 \cdot 10^{-12}\,\Omega\,\mathrm{m}^2$. With increasing $P$, the influence of the film resistance dimishes. One might expect this resistance to be proportional to $1/P$; but a greater rate was observed. The reason for this discrepancy might have been that the films were damaged at the greater loads, making clean carbon contact spots possible.

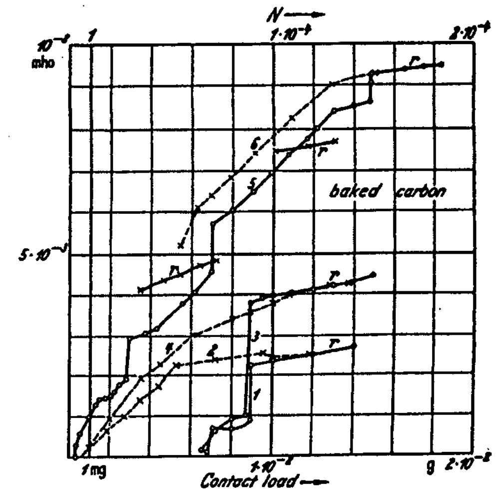


Fig. (9.03). Conductance $1/R$ of contacts of baked carbon, plotted against the load $P$. Hardness $H = 3.7 \cdot 10^8$ N/m². The reversible branches which have been run through several times, are marked with $r$. They have a smaller slope than the other solid lines, which are virgin curves for increasing $P$. Irreversible curves for decreasing $P$ are dashed

The film on the carbon increases the slope of the $r$-branches in Fig. (9.04) to some degree; this might have a slight influence on the sensitivity of microphones, but by no means is decisive for the microphone effect, as was once believed.

At elevated temperatures (500 to 600 °C) an oxide, say a tarnish, can be generated on the carbon

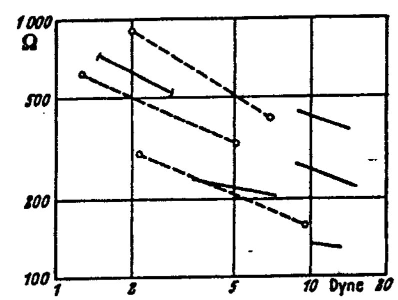


Fig. (9.04). Reversible branches (the dashed ones belong to GOUCHER) of $RP$-characteristics of microphone carbon contacts

[1] GOUCHER [1], cf. CHRISTENSEN and PEARSON [1].

high load, the current is unable to heat the total area to melting during the pulse. The observation is typical for variation of $P$ at a constant current.

*Resistance welding*[1] is the only deliberate welding method with high temperature that is related to the contact theory as it is treated in this book. For details of the methods, we refer to Welding Handbook[2] chapter 30. We add interesting conclusions from application of the $\varphi\vartheta$-relation and of Fig. (21.04). A typical arrangement is shown in Fig. (29.08) where $b, b$ are the plates to be welded together and $E, E$ the water-cooled electrodes which produce pressure and conduct the current. They are softer than the plates in order to provide good contact.

In the beginning there is metallic contact only in very small spots, giving a contact resistance favorable for rapid heating. After a millisecond, the resistance in the weld-surface will have reached its minimum. Further heating is required in order to obtain the greatest tensile strength of the weld.

During the whole welding process the original contact spot remains the warmest section, not only because it offers the greatest resistance in the circuit, but still more because this section is most distant from the cooled electrodes.

The high temperature entails a high resistivity that is favorable for concentration of the heating to the weld.

Table (29.09). *Spot-welding of metal plates*

| Welding contact | Contact load N | Amplitude of contact voltage $U$ V: between electrode and plate | Amplitude of contact voltage $U$ V: between the plates | Amplitude of current A |
|---|---|---|---|---|
| Fe plates | 700 | 0.7 | 0.3 | 5000 |
| Al plates | 700 | 0.3 | 0.1 | 13000 |

A thorough calculation of the temperature distribution is obstructed because no temperature equilibrium is reached during welding. There is also insufficient space for long constrictions to develop. Therefore, the temperature cannot be computed from the voltage using the $\varphi\vartheta$-relation. Typical a. c. amplitudes of voltages during welding are given in Table (29.09).

## § 30. About stationary contacts in practice[3]

**A. Introduction.** While the preceding chapters deal with the physics of stationary contacts seen from a theoretical aspect, this chapter gives examples of how to use the knowledge for understanding stationary contacts as they appear in practice. Here the word stationary expresses that switching during applied voltage appears either seldom or not at all and that the *wear*, which occurs at switching current by the contact, need *not* be *considered*. Typical examples of a stationary contact in the sense of this chapter are bolted bus bars, the terminal contacts of an ammeter, even the plug-socket contact, and also many relay contacts.

[1] Survey: Dixon and Taylor [1].
[2] Published in 1963 by American Welding Soc., New York.
[3] A completion of this Chapter as to moving contacts is given in § 70.



The word practice in the caption of this chapter expresses that the main concern is how to avoid obstructions of the contact operation, for instance by alien particles, films, etc. We shall even base the definition of *contact types* on the kinds of obstruction they are especially subjected to. An appropriate subdivision must apply not only to static contacts but also to contacts which switch during current flow, especially since the difference between "static" and "switching" contacts is sometimes vague.

We distinguish

I. *Permanent contacts* as clamped bus bars and the solderless wrapped contacts of Bell Telephone mentioned on p. 159.

II. *High load* ($P$ = several 10 N) *interruptible* contacts as terminal contacts of ammeters, plug and socket, knife switches.

III. *Medium load* contacts particularly in air, $P = 0.3$ to 1 N, as in intercommunication switches.

IV. *Low load* contacts with $P$ around 0,3 N.

V. *Microcontacts* with $P < 0.2$ N and open contact voltage $< 0.05$ V.

The wear and other phenomena connected with the moving of contacts is treated in Parts III and IV of this book.

**B. Type I, permament contacts**, cf. Section D. The reliability of these contacts is always due to cold welded spots produced at contact make and then protected by the load against (transient) openings, which otherwise could lead to breathing and loss of the weld, see § 22 G.

Into a base metal contact, that has breathed, oxygen diffuses and oxide continues to develop. The resulting rise of the contact resistance is often evident from increasing contact temperature. Loss of load in clamped contacts because of creep may lead to elastic peeling of welds, cf. Fig. (28.06). Therefore, it is often advisable to clamp with spring washers.

A means to protect aluminum bus bars with their relatively great contact spots which are sensitive to even very short breathing is to silver-plate the overlapping portions of the bars. Connor and Wilson [1] describe experiences with silver-coated joints of aluminum conductors. They and also Burley [1] and P. Quinn [1] discuss means

for the plating. Good platings showed a greater tensile strength than the basic metals Al and Ag.

When "permanent contacts" are constructed so that they do not breathe, their life is practically infinite. RICHTER and SCHADE [2] observed that clamped Al-contacts which were good the first year remained fit for use 20 years more, at least. MASON [1] computes that Bell Telephone's wrapped contacts will well have a life of 40 years. His conclusion is supported by MILLS [1].

A *remark* about *plastic flow* in a nominally *flat contact*. If the contact members approach perfect smoothness, plastic flow would require a pressure about equal to the yield point in the total area. The much smaller apparent pressure applied in practice would not be able to rupture the tarnish film. That normal a-spots with metallic portions develop is due to the actual waviness and asperity of the surfaces. Fig. 5 in WILLIAMSON [2] illustrates this fact by showing that low contact resistances (little varying with the roughness) appear at roughness amplitudes above $0.5\,\mu m$; but with smoother surfaces, tarnish films remain because the resistance rapidly increases with decreasing roughness amplitude. One sees that very smooth contacts can be unwelcome. Metallic contact members can be adapted to each other by repeated make and break of the contact, without current, and in the same position. This results in flattening the surfaces. At further contact make, the members meet in points with a relatively small curvature, and the contact spots may be produced by merely elastic yielding. In this state, their films endure the deformation without rupturing; they add their resistance to the constriction resistance and are able to prevent any weld.

A *note on the effect* of the thickness of alien films in the formation of fissures. Fissures probably appear as easily in thick alien films as in thin ones provided the basic metal yield plastically. Does this imply that the thickness of the film has no influence on the final contact resistance? The most elementary experience contradicts this supposition. It is self-evident that a good contact is not secured by fissures alone, but that metal from the members must be squeezed into the fissures. This surely is more difficult, the thicker the film. In very thick films, the fissures are merely channels into which only air can enter, promoting further oxidation.

**C. Separable contacts of high mechanical load.** Plug-socket contacts often have visible oxide films which would insulate if undamaged at make. The actual conduction is produced either by mechanical rupture of the film at closure or by fritting. The available voltage of 100 V or more guarentees frittings. The final adaptation of the a-spots may leave the contact voltage at the order of 0.1 V. The contact is correspondingly heated rendering the socket warm to the touch. This, however, can be tolerated.



Terminal contacts of ammeters, etc. must have small resistances, i. e., their contacts must have metallic or quasi-metallic a-spots in both members. The clamps are made of brass and are usually slightly greased in order to prevent oxidation. A load of several hundred Newtons is reached; cf. Table (30.03). This load and a twist secure a sufficiently small resistance between clamp and copper wire.

Plug and socket connections in Telephone Selectors with a load of 1 to 3 N and voltages of the order of 5 V provide insufficient means to break tarnish films. Palladium plated, they function satisfactory; cf. Section E.

**D. Resistance measurements of clamped contacts**[1] **with force produced by a screw.** The contact between a conducting wire and a terminal screw is suitably investigated in a cricuit sketched in Fig. (30.01) in which the resistance of the wire is not contained in the measured resistance. A sufficiently exact calculation of the screw pressure can be made according to Eq. (30.02). $M$ in Nm denotes the torque, $h$ m the pitch and $2r$ m the diameter of the thread[2] and $f_1$ and $f_2$ the coefficients of friction in the thread and screw head respectively. In one revolution, the torque does the work $2\pi M$ which suffices for: 1. The displacement $h$ m of the head (equal to the compression of the wire) corresponding to the work $hP$, where $P$ is the contact load; 2. the work of friction in the thread of about $2\pi r_1 P f_1$; and 3. the friction work between the head and the compressed wire $2\pi r_2 P f_2$, where $r_1$ and $r_2$ are average values of the axial distance of the contact spots at the thread and head respectively. The contact load in both the thread and the wire is assumed to be constant during one twist. If during this time $P$ varies, one may consider a sufficiently small part of one turn whereby the factor introduced finally cancels. Since $r_1$ is somewhat smaller, and $r_2$ somewhat greater than the radius of the screw, say $r_1 = r/1.1$ and $r_2 = 1.1\,r$, we finally obtain

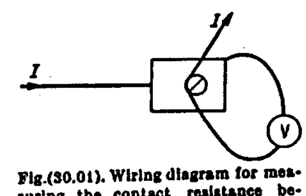


Fig. (30.01). Wiring diagram for measuring the contact resistance between a wire and a terminal screw

$$2\pi M = Ph + 2\pi r P \frac{f_1}{1.1} + 2\pi r P \cdot 1.1 f_2 \quad \text{Nm}$$

or

$$P = \frac{2\pi M}{2\pi r\left(\frac{f_1}{1.1} + 1.1 f_2\right) + h} \quad \text{N} \qquad (30.02)$$

[1] See R. HOLM [29] p. 145, cf. R. HOLM et al. [8] § 12.

[2] American screws are commonly characterized by two numbers. For instance in 8 by 32 the number 8 is a conventional reference to a major diameter of 0.1040" and the number 32 means 32 threads per inch. In Germany a screw with a major diameter = 4 mm is called a M4 screw.

This equation is suitable for computations. Substitution of conventional values shows that friction makes the clamping force about ten times smaller than it would be without friction. Exact precalculations are impossible since $f_1$ and $f_2$ vary considerably from case to case. Friction coefficients as small as 0.15 to 0.25 may occur if some oil adheres to the screw. The coefficients increase the more the oil film is squeezed out of the contact, for instance, by screwing back and forth; e. g., on brass and steel to $f = 0.4$ and on aluminum and zinc to 0.7.

Table 30.03 gives typical examples of the behavior of contacts clamped with screws in normal atmosphere[1] and kept clamped during the measured times.

The contacts were kept at a temperature of 100 °C from the middle of the first month to the end of the second (with the exception of some days when measurements were made). At the end of the first month they were all drawn tighter at room temperature by the same torque $M$ as was used initially. This was possible because, by creep of the metal of the wire, the contact load had given way to a certain degree[2]. It is strange that as a rule, the resistance did not change on tightening the contacts. However, during the noted time, a tenth of the contacts increased their resistances considerably while 1 % decreased them markedly. It is evident that the contact still adhered in original metallic points, and that new metallic spots were not usually generated on tightening the contact.

Table (30.03). *Extract from measurements on screw contacts carried out in F. L.*

| Test | Wire and its diameter | Clamp ledge | Method of cleaning[3] the wire | Screw diameter | Torque $M$ | $P$ about | Resistance at room temperature $10^{-4}\,\Omega$; $t$ in months after clamping | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | mm | | | mm | Nm | N | $t = 0$ | 0.5 | 2 | 12 |
| 1 | Cu 0.5 | Ni | oil | 2 | 0.13 | 280 | 0.08 | 0.05 | 0.05 | 0.04 |
| 2 | Ag 0.5 | Ni | no cleaning | 2 | 0.057 | 120 | 0.18 | 0.20 | 0.25 | 0.43 |
| 3 | Al 0.28 | Ni | cotton | 2 | 0.057 | 90 | 1.3 | 1.9 | 2.2 | 2.2 |
| 4 | Al 0.28 | Ni | oil | 2 | 0.057 | 110 | 0.27 | 0.30 | 0.28 | 0.13 |
| 5 | Al 0.28 | Zn | oil | 2 | 0.057 | 90 | 0.8 | 18 | 400 | 1600 |
| 6 | Al 0.28 | Ag | no cleaning | 2 | 0.207 | 350 | 0.2 | 0.2 | 0.4 | 1.1 |
| 7 | Al 0.28 | Zn | no cleaning | 2 | 0.207 | 280 | 1.0 | 4.5 | 18 | 23 |
| 8 | Al 0.28 | Zn | oil | 2 | 0.207 | 320 | 1.7 | 3.2 | 0.84 | 0.83 |
| 9 | Al 0.6 | Ni | no cleaning | 4 | 0.207 | 130 | 0.43 | 0.48 | 0.50 | 0.58 |
| 10 | Al 0.6 | Ni | no cleaning | 4 | 0.407 | 200 | 0.14 | 0.15 | 0.15 | 0.15 |
| 11 | Al 0.6 | Ni | oil | 4 | 0.407 | 250 | 0.09 | 0.09 | 0.08 | 0.08 |

[1] R. HOLM and collab. [8] § 12.

[2] Cf. § I.

[3] *Oil* signifies that the contact members, after abrasion with emery in oil, were clamped together in the oily state; and *cotton* means that the members were only rubbed with cotton soaked in alcohol.



*Discussion of Table (30.03).* The importance of the pressure from the screw is remarkable. Really poor contacts appear only at $P \approx 100$ N, while $P \approx 200$ N and more produces lasting conductance. Only zinc contacts form an exception[1]. Consequently it is necessary to choose as great a clamping force as the conducting wire can endure without losing the necessary tensile and bending strength outside the terminal[2]. The thinnest aluminum wire in the table, for instance, did not endure more than $P \approx 300$ N.

Hardly any influence of the increase of temperature is noticeable. This experience is based on experiments with contacts at high temperatures carried out during a year. On copper contacts (cleaned with cotton) at 150 °C and $P = 350$ N, only a decrease of the resistance followed; on aluminum contacts (oil) at 150 °C and $P = 50$ N an increase of the resistance of only about 10% was observed. A varying temperature impairs a contact more than a constant temperature, equal to the crest of the temperature variations.

Among observations on aluminum in Germany we mention those by RICHTER, together with SCHADE[3], and those of GEBAUER [1] and VIEHMANN [1] and [2] should be noted. The first mentioned investigators found that screw contacts which were good during the first year remained fit for use for 20 years more. Rivet and twist contacts on the other hand were unreliable. The crimping method which VIEHMANN recommended for stranded aluminum wire is widely used in pressure-type crimps and also for terminating either solid or stranded copper wire. For instance, the strands are enclosed in a metal sleeve, preferably of copper, on which a lasting pressure is exerted. Since the wires can not escape, strong plastic deformations and cold welding appear in the mutual contacts. Of course, the sleeve material must also yield in the contact spots in order to provide a good connection. Wires of alloyed aluminum that creeps less than the pure material are suitable. Notwithstanding the diminished creep and the cold weld, the use of spring washers or similar elastic joints procuring a lasting pressure is recommended, since a disturbing gradual yielding will always exist.

**E. Medium and low load contacts in air. Printed circuit boards.** In early intercommunication switches, base metal contacts of beryllium-copper, phosphor-bronze, nickel-silver, have been used (many still in use). In Germany, much bronze slider on brass lamellas were used. All were disposed to tarnishing. However, the relatively high load and

[1] Cadmium behaves like zinc.

[2] Experience has shown that a three to ten times greater contact load is needed to give fairly clean, flat contacts the same conducting surfaces as are obtained with wires.

[3] RICHTER [1]; RICHTER and SCHADE [2].

high voltage, about 1 N and about 50 V, sufficed to produce a-spots either mechanically or by frittings.

For better performance, one resorts to noble metals. Silver the cheapest of them, has the disadvantage of tarnishing in a sulphur containing atmosphere. Palladium, the next in price, is resistant to sulphur but it catalyzes the formation of polymers when sliding in atmospheres containing organic vapors; see § 25. As a compromise, alloys of silver with 30 to 50 % palladium are expedient[1].

In modern telephone techniques the requirements have become more stringent as the trend goes toward the use of miniaturized multiple contact connectors and printed circuit boards. Here load and voltage are too low to secure a-spot with base metals. With respect to costs, one is restricted to noble-plated material.

The Bell Telephone Laboratories[2] have conducted experiments with various platings exposed (unmated and sheltered against dust) 1/2 year to industrial and maritime atmospheric evironment. With $P = 0.25$ N and contact voltage below 0.01 V, which is too low to rupture the films mechanically or by frittings, only the following coatings were found to be satisfactorily protective:

A 2.5 µm Au plating and the duplex coating of 0.5 µm Rh over 0.75 µm Ag. Au over Ni, Rh over Ni and Au over Ag (sulfiding environment) are unsatisfactory because Ni and Ag migrate through pores in a thin Au coating.

For higher medium contact forces, the following coatings were protective: 0.5 µm Rh over 7.5 µm Ni, 0.5 µm Rh over 5.0 µm Ag, also 2.5 µm hard gold. The composite coating over Cu requires 7.5 µm Ni because a deforming Ni leads to cracks in Rh. Au-coatings lead readily to galling, and lubrication at least for the wearing-in process is recommended; cf. p. 175.

In recent years, new successful plating baths have been developed; see for instance VINES [1]. Pd-platings are widely accepted when polymers can be prevented. Massive Pd is very soft compared with Pd-platings (300 D.P.N.); 5 µm thick layers of Pd can be made sufficiently porefree and crack-resistant. They give a satisfactory resistance against wear[3].

The structure of the deposits is difficult to control. ANGUS [3] presents the following picture. The growth seems to start from a series of nuclei on the surface of the substrate. The unplated spots give rise to a network of fine submicroscopic pores. Their existence is deduced

[1] See BORCHERT [1].
[2] See U. B. THOMAS [1], [3], and R. G. BAKER [1]. Cf. the valuable review by KEEFER and GUMLEY [1].
[3] Cf. ANGUS [3], [4].



from chemical evidence and from the fact that as the thickness in-
creases beyond 1 µm, the coatings become progressively more protec-
tive. However, when they grow beyond 3 to 5 µm, microscopic cracks
very often develop.

F. Medium and low load contacts in sealed chambers. During World
War II, an attempt was made by the Bell Telephone Laboratories to
"achieve the ultimate in relay contact reliability by operating relays in
sealed containers". It was hoped that sealed containers would protect
against undesirable atmospheric conditions and dust. However, the
result was a failure. Corrosion increased and insulating deposits ap-
peared on the contacts, now known to be polymer deposits mentioned
in § 25 and § 49 B. The deposits originated essentially from vapors of
insulating materials; also solder flux, which is difficult to remove, caused
trouble.

These disturbances are avoided in the modern construction of *reed relays*[1] (in German: Schutzrohr-Relais). The reeds are fused into glass containers where a good vacuum is maintained during the manufacture, and necessary welds are performed by means of the electron beam method without solder. The basic idea of a reed relay is illustrated in Fig. (30.04). The magnetizable reeds[2] have their outer ends fused into the glass tube and overlap one or two mm at the inner ends, separated with a gap of about 0.15 mm. They are magnetized by the operating coil (with 50 to 100 ampere turns) to a flux density of the order of 0.5 Wb/m². The ends then attract each other and make contact in a few ms. When the coil current is interrupted, the elasticity of the reeds snaps open the contact.

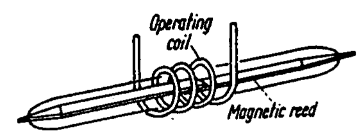

Fig. (30.04). Reed relay

It has been necessary, at least in present manufacturing, to cover
the contacting areas with a metal of no or little tendency to tarnish such
as gold or wolfram. In order to secure a reasonable heat dissipation, the
container is filled with an indifferent gas ($N_2$ plus a few % of $H_2$ at a
pressure of about 1 atm). Nevertheless, the maximum current is only
1 A. The contact producing load is 0.05 to 0.2 N. The life of reed relays
with arc quenching can reach $10^5$ to $10^9$ operations[3]. It is usually limited
by material transfer with interlocking; see § 65.

About constructions with magnetically biased relays, assemblies of relays etc. see HOWGAARD et al. [1] and DUMBAULT [1] p. 80.

[1] See HOWGAARD et al. [1], DUMBAULD [1], WAGAR [1] and WOLAK [1], REUCH [1].
[2] An iron-nickel alloy that can be fused to glass.
[3] OETTL [1] gives an interesting review of the number of operations reached with different contact types.

When considering microcontacts, notice that the range of $P$ places the reed contacts on and below the border between low load and microcontacts.

G. **Microcontacts**, particularly with very low voltage, less than 0.05 V and $P$ smaller than 0.2 N. In the presence of disturbing tarnish, no metallic portion can be generated in the contacts, neither mechanically nor by frittings. The contact surfaces must be initially clean, which means noble metal contacts if a neutral atmosphere is not provided. Frictional polymer deposit must surely be avoided, just as was required for low load contacts; cf. the discussion on polymers in § 25. Here we concentrate our interest on another disturbance, namely on the influence of small particles which may be invisible for the naked eye but which frequently are present, for instance, *air-borne dust particles*.

A microcontact with $P$ below 0.2 N can be kept *open* by such a single particle; cf. WILLIAMSON et al. [1].

Fig. (30.05) represents the analysis given by WILLIAMSON et al. A dust particle is pictured between the contact members $I$ and $II$. In the upper figure, the member $I$ (solid line), considered as having a hard spherical surface, is carried by the area $A$ on the top of particle C, and no contact exists between $I$ and $II$. The dashed line pictures the member $I$ in such a position that it is still partly carried by the particle but has access to a feeble contact at $B$. $AB$ is the radius of the "region of influence" of the particle. Within this region no contact can appear. But with member I moved slightly to the right a full contact would develop at $B$.

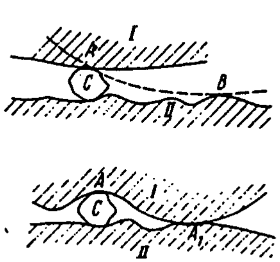

Fig. (30.05). Effect of dust particle in a contact. Vertical scale is exaggerated

So far, it was assumed that the load $P$ has deformed member $I$ (and the particle) only slightly at the contact spot. The area $A$ of the spot is in accordance with

$$P \approx A\,H \qquad (30.06)$$

where $H$ is the hardness of the softer member. Eq. (30.06) holds true provided that the deformation is "deep plastic"; cf. § I. The maximum $A$ will essentially equal the cross section $A_p$ of the particle.

The lower figure pictures the event that the load is able to deform the member $I$ as to produce contact areas $A$ on the particle and $A_1$ with member $II$. Hence

$$P \approx (A + A_1)\,H \qquad (30.07)$$



where now $H$ may refer to member $I$. The particle can not prevent the contact at $A_1$ when the load $P$ is great enough to satisfy (30.07).

We conclude from this discussion that a particle with the cross section $A_p$ is likely to cause open contact if

$$\left.\begin{array}{l} P < A_p H \\ \text{but will not be able to do it if} \\ P > A_p H \end{array}\right\} \qquad (30.08)$$

Fig. 7 in WILLIAMSON et al. [1] shows that Eq. (30.07) is essentially true. In practice $H$ is of the order of $10^9$ N/m² and dust particles[1] have $A_p$ between the limits $2 \cdot 10^{-12}$ and $2 \cdot 10^{-10}$ m², corresponding to about

$$\left.\begin{array}{l} 2 \cdot 10^{-3} < P < 0.2\ \mathrm{N} \\ \text{or in other units} \\ 0.2 < P < 20\ \mathrm{g} \end{array}\right\} \qquad (30.09)$$

Meaning: with $P > 20$ g, a contact is not likely to be kept open by a single dust particle, but with $P < 20$ g this can happen. That is the reason why we let $P = 20$ g differentiate the contact classes for micro- and low-medium load. The significance of $P = 20$ to 30 g for the surety of contact make has been known for a long time; cf. for example R. HOLM [37] p. 392.

It might be questioned whether a special name should be given to contacts with extremely low load; for instance when $P$ = a fraction of a gram. This seems not to be advisable because all microcontacts require noble metals or vacuum or an indifferent atmosphere independent of $P$ in the range below 0.2 N. Only the maximum current must be adapted to the load so that the contact voltage remains below the softening limit. Otherwise considerable deformation could be produced, jeopardizing the adjustment of the contact.

We now pass to particles that do not come from the air. CHAIKIN [1], [3] has evidenced that during machining and incorporation of the parts in the relay, base material particles can be transferred from the tools to the contact members (cf. Fig. [28.06]). He uses sensitized test paper pressed against the surface concerned to show the presence of the transferred particles; see for example Fig. 1 in CHAIKIN [1].

The insulating power of such deposite should not be tested by sliding the wire probe[2] along the surface because it can easily gather *several* disturbing particles and experience a higher insulating power than could derive from one particle.

[1] R. J. HAMILTON [1], examples also in PFISTERER et al. [1].
[2] The probe is described in § 9C.

When small particles are very numerous, the discussion referring to Fig. (30.05) may fail to describe their influence on contacts because many particles may cover the load bearing area and together carry the load. The tumble-polish that is connected with electroplating seems likely to produce this kind of contamination; see, for instance, LISH [1].

**II. Stored lubricated contacts.** If open contacts of base metals shall be stored for a long time, and must be available for use at any moment, it is customary to protect them with a lubricant. Of course, the lubricant must not contain, or in the course of time, form ingredients which corrode the metal. Thick films of the lubricant protect somewhat better than thin ones but have the tendency to gather dust and form a disturbant mash. This disadvantage is avoided when a *monolayer* of the lubricant is used. Still the protection is considerable as shown in Table (30.10) based on measurements by R. HOLM et al. see R. HOLM [37], p. 163 and 164.

The samples, initially practically clean, were tested against a gold mating member with $P = 1$ N without impact or sliding, in order not to damage the tarnish. The monolayer (sometimes called epilamen) was placed on the sample in the following way. One drop of a very dilute solution of a long-chain polarized compound in trichlorethylene, carbontetrachloride, or a similar solvent was put on the specimen and then wiped off. Afterdrying, a monolayer remains and sticks firmly to the sample. The tested epilamen[1] was a compound of a long-chain paraffinic hydroxan acid with straight-chain paraffinic hydrocarbon, giving the epilamen a particularly dense structure. It is seen that the final contact resistances are much greater without than with protection, which means that the tarnish grew relatively slowly under the epilamen. However, testing the samples *against* members of the *same metal*, showed practically *no change* during the time of the experiment. This means that all tarnish concerned was very vulnerable. The table contains data averages. The spread of the data was great, with a standard deviation of about 50%.

Table (30.10)

| Material against Au | Resistance in $10^{-3}\,\Omega$ initial | without protection after 2 weeks | without protection after 1/2 year | with epilamen after 1/2 year |
|---|---|---|---|---|
| Ag | 1 | 3 | 20 | 3 |
| Cu | 10 | 100 | >1000 | 10 |
| Ni | 10 | 100 | 200 | 15 |
| W | 3000 | 5000 | 6000 | |

[1] The epilamen was manufactured in the organic laboratory of Siemens & Halske A. G. Berlin, Germany.



With an epilamen formed of stearic acid, a resistance increase similar to that noted for epilamen-covered contacts after 1/2 year in Table (30.10) appeared within 2 months. Thus different lubricants have different protecting power. This has been stated also by CHIARENZELLI et al. [1] operating with thin lubricating films. Whether they were monomolecular or not was not discussed. These authors found a mixture of paraffin and syntetic oil to be a promising protector. They made the tests with sliding contacts.

**I. Contacts in measuring apparatus[1].** In this context, it will be appropriate to mention *lever resistance contacts* (Kurbelkontakte) and *wire rheostats*, although they have certain qualities which would motivate their treatment in Part III of this book. For obvious reasons the resistance material can not be of noble metals. The good conduction in the contact has to be secured by the slider scratching the tarnish film, if the device is used in air without a protecting lubricant. If a very low contact resistance is required, greasing is recommended. Lubrication should be renewed once a month because the grease deteriorates in the air; perhaps partly because of catalytic action of the metal.

The wear of such contacts is of interest. Typical data are given in Table (30.11).

Table (30.11). *Typical numbers for sliding contacts in measuring apparatus.* Both wear, expressed by[2] $Z$, and hardness $H$ in $10^8$ N/m$^2$ refer to the softer member

| Rotary switch: slider | Rotary switch: stationary contact | $P$ N | $R$ $10^{-3}\,\Omega$ | $Z$ | |
|---|---|---|---|---|---|
| Bronze or silver $H = 8$ | Brass or silver $H = 10$ | 10 | 0.3 | 0.8 | greased |
| $H = 8$ | $H = 10$ | 10 | 0.3 | 16 | dry |
| **Wire rheostat: slider** | **wire** | | | | |
| Cu, 2.5% Be $H = 25$ to 33 | Constantan $H = 14$ | 1 | 30 | 6 | slightly greased |

Terminal contacts for instruments need rubbing closure, which means a certain wear. Wear shall be of the small, not of the micro-type, nearly of the order noted in Table (30.11); cf. BAYER et al. [1].

**J. Contact assembly.** The method of contact assembly is an important factor in the final performance of electric contacts. It would require a large chapter to give a survey of the many methods of riveting, brazing, welding, etc. However, these are manfacturing methods and lie beyond the scope of this book that is devoted to explain contact phenomena. We emphasize the importance of the contact assembly,

[1] Cf. R. HOLM [34]. [2] As to $Z$ see Eq. (41.03).

metal may be blown out – sputtering – during the intense boiling in the arc spots[1]. Coherence of the refractory diminishes the sputtering. It is common to use silver compounds in air blast breakers and sintered copper compounds in oil breakers. The idea is that the silver will resist oxidation in the air[2], which the copper does not. Unfortunately, less is gained than expected since silver together with wolfram or molybdenum oxidizes to silver-wolframate or -molybdenate[3], glassy slugs which are able to produce insulation at contact make.

**G. Sliding contacts for resistors and apparatus**[4]. Different hardness in rider and stationary member is used in order to minimize wear. In many cases it is particularly important to prevent wear of the stationary member, for instance, on film potentiometers where a relatively soft gold wire may slide on the harder metal deposited on glass. For some servo-mechanisms relatively soft brushes of metal-graphite type may be used on silver or silver-alloy rings.

Wire-wound potentiometers usually have wires of oxidizing materials, as manganin, constantan or chromium-nickel-alloys, on which a somewhat harder rider slides, consisting of bronze, nickel or hard platinum alloy. The rider must be able to abrade thin oxide films. Oxidation is kept low by lubrication, while for purposes of high precision the whole potentiometer may be kept submerged in oil.

[1] W. R. Wilson [2].
[2] Oxidation of silver occurs above 550 °C.
[3] Keil [4] p. 223.
[4] Cf. § 30 I.

# Appendices

## Appendix I. Elasticity, Plasticity and Hardness

**A. Introduction.** The concepts of elasticity, plasticity and hardness are discussed with respect to their application in the theory of contacts, viz. for the determination of actual contact areas. Simplification is gained by reference to more complete but easily read treatments by Tabor [*1*], Kuhlmann-Wilsdorf [*1*], and Bowden and Tabor [*12*], Chapter XVI.

Two ideally hard bodies cannot touch in more than three points. In these contact points the pressure $= \frac{\text{mechanical load}}{\text{area}}$ would be infinite. But in reality, the material yields and thereby defines contact areas. Additional contact spots may then be generated. The yielding is *elastic* if the initial shape of the bodies is restored when the stress is removed, and *plastic* if a deformation remains. It is essential to note that a plastic deformation is always surrounded by an elastic deformation which delivers the load carrying counter-force.

The mathematics for this phenomenon can be worked out only approximately. The most elaborate calculations concern *ideally* elastic or plastic contact members with simple shapes, particularly in the cases where one or both members are spherical or both are cylinders making a cross-rod contact. Also, perfect isotropy is assumed, with Hooke's law valid for the elasticity, and no strain hardening. We can then extend the discussion to nominally flat members whose actual surface roughness can be represented by spherically curved protuberances. The time dependent phenomena will be treated in Section H.

Certainly, no crystal has spherical symmetry. However, contact materials usually are polycrystalline and behave essentially isotropically in as much as the dimensions of the crystallites are much smaller than any curvature radius concerned.

**B. Hertz' formulas for ideally elastic indentations.** Hertz' formulas (Hertz [*1*]) with many additions are cited in Roark [*1*] p. 287 ff. They are repeated here to the extent they are used in the present book. Between two spheres the contact surface is practically flat and circular

with the radius $a$, given by Eq. (I,1). Eq. (I,1) gives the general formula for different members (1) and (2), with $r$ = radius of curvature (positive for a convex and negative for a concave surface), $\mu$ = POISSON ratio between lateral and longitudinal strain under the condition of longitudinal stress, $E$ = YOUNG's modulus of elasticity and $P$ = load. For $a \ll r_1$ and $r_2$ (cf. below the reference to STOREY)

$$a = \sqrt[3]{\frac{3}{4} P \left(\frac{1-\mu_1^2}{E_1} + \frac{1-\mu_2^2}{E_2}\right)\left(\frac{1}{r_1} + \frac{1}{r_2}\right)^{-1}} \qquad (I,1)$$

giving Eq. (I,2) with $r_1 = r$ and $r_2 = \infty$, $E_1 = E_2 = E$ and $\mu_1 = \mu_2 = 0.3$, i. e., for a ball against a semi-infinite plane body of the same material with POISSON ratio[1] of 0.3.

$$a = 1.1 \sqrt[3]{\frac{P}{E} r} \qquad (I,2)$$

Here, in a point of contact at a distance $x$ from the center of the load bearing area $A_b$, the pressure is

$$p = \frac{1.5\,P}{\pi a^3} \sqrt{a^2 - x^2} \qquad (I,3)$$

In the center, the pressure has its maximum

$$p_{max} = \frac{1.5}{\pi a^2} P \qquad (I,4)$$

With an elastic indentation producing $A_b = \pi a^2$, the centers of the spheres approach each other by

$$y = a^2 \left(\frac{1}{r_1} + \frac{1}{r_2}\right) \qquad (I,5)$$

where $y/2$ is the sum of the indentations and the other half of $y$ is due to elastic deformation in the bulk of the spheres. When not complete spheres but only spherical humps produce the indentations, the elastic deformation in the bulk is smaller than $y/2$. The distance $y$ is often labeled "compliance". Eq. (I,5) is also valid for negative $r$.

For a contact between a sphere and a plate, the plate is required to be at least $6y$ thick; otherwise $a$ deviates considerably from (I,1).

The formulas (I,2 to I,5) are also *valid* for the contact between two *crossed cylindrical rods* with the same $r$ and $E$, and $\mu = 0.3$ for both members.

The more complicated formulas for the elliptical contact area of crossed rods of different radii and materials are given in ROARK [*1*] p. 289. He also treats the contact between parallel cylindrical rods, a case which is difficult to realize in practice.

[1] Actual values of $\mu$: 0.28 for Fe; 0.36 for Cu and Ag; 0.39 for Pt.



According to SHTAERMAN (see C. STOREY [*1*]), Eq. (I,2) is theoretically valid up to considerable elastic deformations, even beyond the actual limits for elastic deformation. For instance, when (I,2) gives $a = 0.4r$, the complete formula gives $a = 0.396\,r$.

Eq. (I,6) gives the time, $t_i$, of a perfectly elastic impact[1] of a spherical indenter (radius $r$ and mass $m$) hitting a heavy flat anvil with the velocity $v$ m/sec:

$$t_i = 2.74 \left(\frac{m^2}{v r}\right)^{0.2} \left(\frac{1}{E_1} + \frac{1}{E_2}\right)^{0.4} \text{sec} \qquad (I,6)$$

where $E_1$ and $E_2$ are the YOUNG's moduli of elasticity.

**C. Plastic deformations. Dislocations.** Plastic deformation of crystalline solids proceeds by slips (in German Gleitung), and the understanding of slips is based on the theory of *dislocations* (in German Versetzung) which will now be sketched. At least in rough outlines, these facts are rather commonplace. This includes the fact that any mechanical strength (i. e., the limit where the deformation passes from elastic to plastic) such as tensile strength and hardness is ultimately shearing. It is equally well known that slips follow along crystallographic planes and are essentially directed with a 45° inclination against the direction of principal stress. Reasonable calculations indicate that the stiffness of an ideal lattice against the starting of slip should be hundreds of times higher than the observed shear strength. Thus, the actual behavior of slips proves that lattices are not ideal. The essential irregularities in question are dislocations. The simplest type of a dislocation (an edge dislocation) is illustrated in Fig. (I,7) which pictures a section through the boundary between two atomic layers. Within the region $xy$, the lower layer has one row of atoms more than the upper one. The lateral extension of the irregularity is usually so small compared with the surrounding regular lattice that it is reasonable to speak of a dislocation line being perpendicular to the plane of the figure.

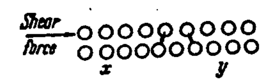

Fig. (I,7). Edge dislocation

Evidently, bonds between upper and lower neighboring atoms are directed to the left on one side and to the right on the other side of the dislocation center as indicated by arrows. A shear force pressing the upper layer to the right is aided by the right side bonds. This results in a smaller stress being required to move such a dislocation than to shear an ideal crystal. This is true even though the dislocation is somewhat anchored at its ends. The edge dislocation moves perpendicular to its axis. When it finally reaches the crystal boundary, the effect is a slide of a cross section of the crystal by the length of one spacing

[1] Eq. (I,6) has been deduced by TABOR [*1*] p. 131 using formulas by HERTZ. For equal spheres the coefficient is ≈ 3.

(or one Burgers vector). Certainly, a macroscopic slip requires the increments of many dislocations.

The easy move of dislocations is impeded by their anchoring in lattice irregularities such as vacancies, interstitial atoms or rather clusters of them, crystal boundaries, and intersections with crossing dislocation lines. *Strain-*(work-) *hardening* results when abnormally numerous anchoring obstacles are produced, for instance, by subdivision of crystallites.

It is understandable that a material can appear with strengths as different as strain hardened and annealed copper. It is rather surprising that a common quality can be characterized within small standard deviations to make reasonable the presentation of normal hardness values (in Table X,1). A plot of such numbers vs the strength of cohesion bonds, as judged from evaporation energies, suggests that the dislocation-anchoring forces are roughly proportional to the strength of those bonds.

One of the reasons for the existence of a normality is that dislocations are impeded from gathering in very tight groups by the action of parallel and equally directed dislocations repelling each other. The repulsion is connected with the fact that adding two dislocations would double the Burgers vector and increase the total dislocation energy. This is proportional to the square of the Burgers vector. Usually the distortion energy in a dislocation is about 2 eV per length of 1 Å. The normal distance between dislocations is some 1000 Å. However, distances down to 100 Å have been recorded and the most regular Ge crystal produced had a dislocation spacing of about 1 mm.

Dislocation lines are made visible in the microscope by the so-called decoration method, obtained by alloying the material with a substance that gathers in dislocations and colors them. In an etched surface, mouths of dislocation lines appear as small pits; for illustrations see Kuhlmann [*1*] Figs. (4) and (14).

Evidently, dislocations will move along (001), (111) etc. crystallographic surfaces in single and polycrystalline bodies. This means that the plasticity can be approximately "ideal" only if there are slip surfaces in all directions. For example, in face-centered cubic crystals there are altogether 12 slip systems. In contrast to this, only the (001) plane type is a slip surface in graphite. For *generation of dislocations* see Kuhlmann [*1*].

**D. Mathematics of plastic yielding.** It has been emphasized that plastic yielding is the result of shearing. A hydrostatic pressure does not produce any plastic deformation. For our purposes, these facts are expressed in v. Mises' Eq. (I,8) with sufficient exactness:

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2\,Y^2 \qquad \text{(I,8)}$$



It gives the criterion for plastic deformation by expressing that plastic deformation will occur in an isotropic material when the principal stresses $\sigma_1, \sigma_2, \sigma_3$ reach values which satisfy (I,8).

From the principal shear stresses

$$S_{12} = \frac{1}{2}(\sigma_1 - \sigma_2); \quad S_{23} = \frac{1}{2}(\sigma_2 - \sigma_3); \quad S_{31} = \frac{1}{2}(\sigma_3 - \sigma_1)$$

it follows that Eq. (I,8) requires a certain amount of total shear stress.

Eq. (I,8) can not be satisfied by a hydrostatic pressure; i.e., by $\sigma_1 = \sigma_2 = \sigma_3$ which would make the left side equal to zero.

The meaning of $Y$ manifests itself when considering the tensile or compression test of a cylinder. At the start of yielding, the axial stress is $\sigma_1$ whereas $\sigma_2 = \sigma_3 = 0$. From Eq. (I,8) it then follows that $\sigma_1 = Y$; consequently, $Y$ is the yield stress, *yield point*. $Y$ would be a constant in absence of strain hardening. However, if the material hardens during the deformation, $Y$ increases correspondingly. It can also happen that the material looses its symmetry. Both events would effect an equation more complicated than (I,8).

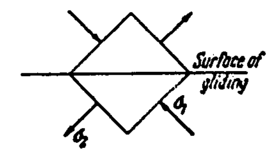


Fig. (I,9). Principal stresses during pure shearing

According to Eq. (I,8) also the shear strength would be $Y$.

To prove this, consider pure shearing, for simplicity without a hydrostatic pressure[1], as illustrated in Fig. (I,9). Evidently $\sigma_1 = -\sigma_2$ and $\sigma_3 = 0$. Inserting in Eq. (I,8) we find $\sigma_1 = |\sigma_2| = Y$. The shear force in the diagonal area of the elementary cube of the figure is $\frac{1}{\sqrt{2}}(\sigma_1 + |\sigma_2|) = \sqrt{2}\,Y$ giving the shear strength $= \sigma_1 = Y$.

This does not quite agree with measurements, according to which the shear strength is 20 to 30% smaller than the tensile strength[2].

**E. Indentation in an isotropic semi-infinite body, produced by a spherical indenter (ball).** Fig. (I,10) illustrates the indentation after

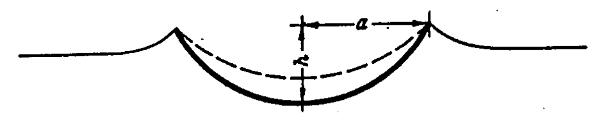


Fig. (I,10). Ball indentation with $a$ = radius of the mouth area and $h$ = depth

plastic yielding. The heavy line represents the indentation before removal of the ball. The finally remaining deformation of the surface is

[1] According to Eq. (I,8) a hydrostatic pressure has no influence on the shear force. It adds the same quantity to any stress $\sigma_1$, $\sigma_2$ and $\sigma_3$ which cancels out in Eq. (I,8).

[2] See Am. Inst. of Physics Handbook (1963) p. 2-62 to 2-68. Cf. § 36C.

indicated by the dashed line[1]. The space between these two lines corresponds to the elastically deformed portion of the indentation. The stress is most inhomogeneous at the rim, and plastic deformation in the contact surface begins close to the rim[2]. That is why the mouth of the indentation remains visible and can be measured even though the load was insufficient to produce plastic deformation overall in the load bearing area $A_b$.

Fig. (I,11) sketches the slip lines under the ball[3] for the case when the plastic deformation is either still in progress or just at the end, finally reaching the broken line at a depth of about $1/2a$. The pressure at the ball surface is not uniform, but distributed according to the values marked; cf. TABOR [1] Fig. 27.

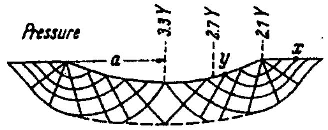

Fig. (I,11). Slip lines in an indentation, and pressures vertical to the wall of the indentation

Imagine that we follow the slip-line $xy$ in Fig. (I,11) and note the changes in the vertical principal pressure $p$ along it. At $x$ this pressure is zero. The horizontal pressures at $x$, which are responsible for the slip, are obtained from Eq. (I,8) to be $p_1 = Y$. The horizontal pressures vary slowly along $xy$ and still are of the order of $Y$ at point $y$. There the plastic deformation requires $p > p_1$, namely at least satisfying $2(p - Y)^2 = 2Y^2$ or $p = 2Y$.

This roughly simplified analysis suggests why (in the ball indentation test) the average pressure against the indenter turns out to be larger than the yield point $Y$. Actually, in the event of a specific depth[4] of about 0.05, $\bar{p}$ is not only $2Y$ but about $3Y$; cf. HILL [1] chapter IX, also TABOR [1] pp. 37, 55, 73 and 104.

**F. The ball and pyramid indentation tests. Hardness.** *The ball indentation test*; see Fig. (I,10). A ball with the radius $r$ is pressed against a flat sample producing an indentation. To begin with, regard the ball as infinitely hard. Then $r$ is the radius of curvature of the indentation. Call $a$ the radius of the circle defining the mouth of the indentation, $A$ the area of the mouth, and $S$ the area of the curved surface of the indentation; $dA$ and $dS$ are elements of $A$ and $S$. The normal to $dS$ makes an angle $\alpha$ with the direction of the load $P$. If the pressure force per-

[1] About the effect of strain hardening on the indentation, see BOWDEN and TABOR [12] p. 336.
[2] The plastic deformation can not begin exactly at the rim when still $p = 0$ according to Eq. (I,3), cf. Fig. (I,11).
[3] under the condition that friction between the ball and the base member does not produce considerable horizontal forces.
[4] See (I,12).



pendicular to $dS$ is $p\,dS$, its component in the direction of $P$ is

$$p\,dS\cos\alpha = p\,dA$$

Integration gives

$$P = \int p\,dA = \bar{p}A$$

where $\bar{p}$ is a kind of average of the pressure against the wall of the indentation. Hence, since $A = \pi a^2$

$$\bar{p} \approx \frac{P}{\pi a^2}$$

Let the depth of the indentation be $h$; see Fig. (I,10). We introduce the *specific depth* $h/r$ as a dimensionless characteristic of the geometry of the indentation and label it $D$. It is readily seen that

$$D = \frac{h}{r} \approx \frac{1}{2}\left(\frac{a}{r}\right)^2 \qquad (I,12)$$

showing that $D$ can be calculated without measuring $h$.

It is apparent that indentations with the same $D$ are geometrically similar and have the same degree of deformation (as defined by [I,8]) in homologous points[1]. Notice that Eq. (I,8) does not contain dimensions.

Actual balls are not perfectly hard and therefore $r$ in Eq. (I,12) should be corrected as shown in E. HOLM et al. [5] p. 231.

Typical measurements with the average pressure $p$ plotted against the specific depth $D$ are shown in Figs. (I,13) and (I,14)[2]. The deformation is purely elastic up to the yield point $\bar{p} = Y$, see Fig. (I,13). The corresponding small variation of $D$ is not marked on this figure. All points refer to a visible indentation mouth, thus to plastic yielding at least at the rim of the indentation. The pressure, $\bar{p}$, increases steadily with $D$ and does not have any strict maximum; but after $D$ has reached a value of about 0.03, the further increase in $\bar{p}$ is small.

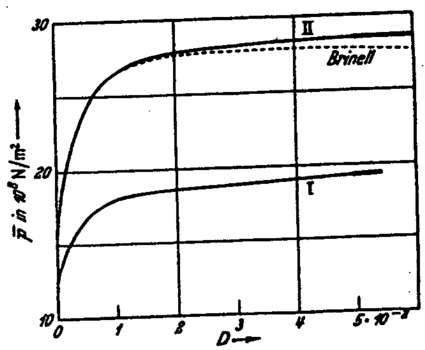

Fig. (I,13). Due to E. HOLM, R. HOLM and SHOBERT [5]. Average pressure $\bar{p}$ plotted again specific depth; $\bar{p}$ along the thick lines is called contact hardness. Steel, strain hardened: I, slightly annealed; II; not annealed

[1] This is an expression of MEYER's [1] similarity law.
[2] Cf. TABOR [1] Fig. 34.

With respect to the moderate accuracy attainable in calculations about contacts, $\bar{p}$ in the range of $D > 0.03$ may be regarded as defining the hardness of the material. The definition must be completed by the stipulation that the indentation pressure shall last about 20 sec

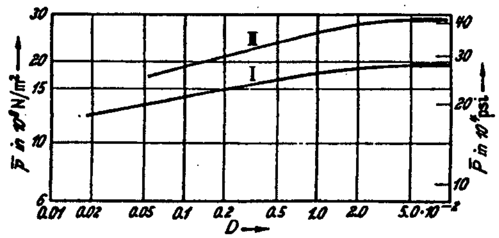


Fig. (I,14). Same measurements as for Fig. (I,13) but plotted in logarithmic co-ordinates

because this time is needed for completion of the (initially relatively rapid) creep. Other definitions of hardness attain a seemingly unequivocal character by referring to the curved indentation $S$ instead of the mouth $\pi a^2$; see below. In order to distinguish the definition used in this book, we speak of *contact hardness* or MEYER's hardness defined by Eq. (I,15).

$$H = \bar{p} = \frac{P}{\pi a^2} \tag{I,15}$$

This requires that contact is made without impact, impression time is 20 sec, and $D > 0.03$.

The mouth area $\pi a^2$ is practically equivalent to the load bearing area $A_b$. Thus, for any shape of $A_b$, it is reasonable to put

$$P = H A_b \tag{I,16}$$

provided it can be assumed that plastic deformation is reached overall in $A_b$.

The load bearing subareas of a nominally flat contact are indentations of different shapes. Types are illustrated in Figs. (36.01) and (36.02). Some of these subareas may be rather deep indentations with $D > 0.03$; others will be shallow with smaller $D$'s (even with no remaining indentation). Consequently, $\bar{p}$ will be smaller than $H$. This situation is expressed by

$$P = \xi H A_b \quad \text{or} \quad \frac{\bar{p}}{H} = \xi \tag{I,17}$$

According to the second expression of (I,17), $\xi$ as a function of $D$ is represented by a curve similar to those of Fig. (I,13) if the ordinate scale marks $\xi = 1$ where $\bar{p} = H$.





As remarked at the end of Section E, the indentation hardness is greater than the yield point $Y$, normally $H = 3Y$ if no strain hardness is produced during making the indentation; otherwise $H$ is up to $4Y$ or even higher. Recalling that $Y$ = tensile strength $S$, we have

$$H = 3Y \text{ to } 4Y \quad \text{with} \quad Y = S \tag{I,18}$$

At the beginning of plastic deformation (in the ball indentation test):

$$\bar{p} = Y \quad \text{or} \quad \xi = \frac{1}{3} \text{ to } \frac{1}{4} \tag{I,19}$$

*Remark concerning the duration* of the load $P$ on an indenter: Immediately after load application, the area of indentation begins to increase by creep. The time until a practically final area is reached depends on the metal. The time is short with steel and relatively long with soft contact metals. At room temperature, $\bar{p}$ may decrease 6 to 14% in 20 to 60 sec corresponding to a decrease of the constriction resistance of 3 to 7%. At the temperature of liquid air, no creep is observed[1].

During sliding, the formation time of any momentary indentation is very short and creep can not develop. Therefore $\bar{p}$ is larger than for stationary contacts.

In the literature, BRINELL's, VICKERS' or KNOOP's hardness numbers are usually given instead of MEYER's. Fortunately, they are sufficiently near (about 7% below) the MEYER's or contact hardness values to be used directly in Eq. (I,15) when they are expressed in N/m².

BRINELL [1] and [2] used ball indenters (from ball bearings) and defined hardness as

$$H_B = \frac{P}{S} \tag{I,20}$$

where $S$ is the curved surface area of the indentation (not the mouth). The reference to $S$ is an artifice without a physical meaning. However, it provides a maximum of $\bar{p}$, as seen on Fig. (I,13), and it is this maximum which BRINELL used for his hardness definition. The corresponding $D$ lies between 0.05 and 0.07.

VICKERS' diamond pyramid indenter has the merit of always producing the same specific depth ($D \approx 0.06$) practically independent of the load $P$. VICKERS calculates with Eq. (I,20) and obtains hardness numbers nearly equal to $H_B$ as a consequence of the choice of the angle of 136° between opposite planes of the pyramid. KNOOP uses a pyramid which produces an in dentation that is 7 times longer than wide[2].

[1] See MULHEARN and TABOR [1].
[2] For details concerning VICKERS and KNOOP hardness see TABOR [1] p. 164 and p. 100 respectively.